

# SHAPE SPLINES AND STOCHASTIC SHAPE EVOLUTIONS: A SECOND ORDER POINT OF VIEW

ALAIN TROUVÉ AND FRANÇOIS-XAVIER VIALARD.

## CONTENTS

1. Introduction	1
2. Hamiltonian equations of geodesics for landmark matching	3
3. Spline interpolation on landmark space: a generative growth model	4
4. Existence results and Euler-Lagrange equation for shape spline estimation	8
5. Numerical experiments	13
6. A stochastic shape spline model	17
7. Shape splines on homogeneous space	22
8. Conclusion	24
9. Appendix	24
References	25

**ABSTRACT.** This article presents a new mathematical framework to perform statistical analysis on time-indexed sequences of 2D or 3D shapes. At the core of this statistical analysis is the task of time interpolation of such data. Current models in use can be compared to linear interpolation for one dimensional data. We develop a spline interpolation method which is directly related to cubic splines on a Riemannian manifold. Our strategy consists of introducing a control variable on the Hamiltonian equations of the geodesics. Motivated by statistical modeling of spatiotemporal data, we also design a stochastic model to deal with random shape evolutions. This model is closely related to the spline model since the control variable previously introduced is set as a random force perturbing the evolution.

Although we focus on the finite dimensional case of landmarks, our models can be extended to infinite dimensional shape spaces, and they provide a first step for a non parametric growth model for shapes taking advantage of the widely developed framework of large deformations by diffeomorphisms.

## 1. INTRODUCTION

Mathematical and statistical modeling of shapes has undergone significant development over the past twenty years driven by a wide range of applications in medical imaging. Initially the focus was on the comparison between two shapes also refereed to as registration. Among others, registration and comparison tools derived from a Riemannian point of view on shapes spaces and diffeomorphic transport have been actively developed during the past few years. This framework was used to represent shapes and study the statistical variation of static shapes within a population. An emerging question of interest is now to study time dependent data of shapes (images, landmarks, surfaces or tensors). The basic dataset is then a sequence of shapes indexed by time. For example, a practical application would be the analysis of follow-up studies in brain imaging.

Several attempts models for the variability of longitudinal data have been proposed recently: a parametric model of growth is proposed in [14, 15], which aims to describe the biological evolution as an iteration of random elementary diffeomorphisms, so called GRID. Focusing on image data, statistical estimation of the parameters are performed with the GRID model in [30, 32]. Other attempts are often based on using an initial registration tool (such as geodesics on a group of diffeomorphisms, see [35] for a large overview) to interpolate the time dependent data with piecewise geodesics [9, 22]. In [9], the model is further developed with the introduction of time realignments to allow the

study of an ensemble of longitudinal data and the computation of an averaged space-time evolution.

From the modeling point of view a typical growth evolution is usually smooth in time. However, the piecewise geodesic models underlying current analysis of time dependent shape data can not prevent a loss of regularity at the observation points. Moreover, from a more classical statistical point of view, piecewise linear regression does not provide the best interpolation framework and from a probabilistic point of view the limiting process underlying piecewise geodesic interpolation is more or less a Kunita flow [23] that has a Brownian like time evolution. These remarks suggest that a better model of interpolation in time should be of higher order than piecewise geodesics, and also closely related to a probabilistic model of random evolution of shapes. The question of time interpolation was addressed in [7] with the use of a kernel on the time variable but still with underlying piecewise geodesics on the data. Random evolutions of shapes have been treated for instance in [8, 21], but here again the model is of first order in the sense that the evolution is not smooth in time.

In this work, we present a second order model for deterministic and stochastic shape evolution as a primary step toward further statistical applications. This work will extensively use the Hamiltonian framework that emerged several years ago in the field of Computational Anatomy (CA) to compute the registration of landmarks (points of interest on an image) [13, 19, 26]. This problem of landmark registration is equivalent to finding geodesics on the Riemannian manifold of landmarks. Thus shooting methods as well as gradient descent have been applied to solve this problem [4, 28].

A first natural step to deal with random shape evolution is to study a stochastic perturbation of the Hamiltonian equations of geodesics. If we consider the evolution of landmarks as a physical system of particles, it corresponds to the introduction of a random force on their evolution providing smooth time random perturbation around a mean geodesic trajectory. However, this in turn leads to a new deterministic counterpart when the stochastic term in the evolution of the momentum is replaced by a deterministic control variable. Given a sequence of time-indexed shapes, and optimizing on the control variable, we end up with a framework for smooth time interpolation between shapes. In the finite dimensional case of landmarks, this approach is directly related to splines on a Riemannian manifold [6, 29, 18, 5]. We will present a Hamiltonian approach to derive the equations for the splines in a Hamiltonian setting. It can be related to the work of [17] since they use a control point of view to obtain the Hamiltonian equations but our approach differs from theirs in that we use the Hamiltonian formulation of geodesics on landmark space. Hence there is a straightforward generalization to other Hamiltonian systems. The possible extension of the Hamiltonian equations of geodesics to curves, surfaces (or embeddings of manifolds in  $\mathbb{R}^n$ ) is the key feature to design second order models of evolution on continuous shape spaces [13, 26, 34] but in this work we will concentrate on the finite dimensional case of landmarks. Note however, that contrary to the initial context of air-craft trajectory planning (where the target manifold is the low-dimensional Lie Group  $SE(3)$  - the Special Euclidean group on  $\mathbb{R}^3$ ), for which splines on Riemannian manifold were introduced - shape spaces in Computational Anatomy are more strongly connected with the infinite dimensional case. Moreover, we will consider both the deterministic and the stochastic setting since these are considered as fundamental ingredients towards statistical analysis of longitudinal data.

The plan of the paper is as follows: we first introduce in section 2 the Hamiltonian equations for the case of landmarks, which constitute the central tool in this article. The section 3 is devoted to the spline model with different metrics and section 4 gathers theoretical results about the global existence in time of the second order deterministic spline evolution, the existence of a minimizer for the spline estimation problem and the derivation of the Euler-Lagrange equations. Then in section 5, we present some numerical experiments about shape spline estimation on 2D shape evolution, highlighting some of their main features. In section 6, we turn to problem of stochastic second order spline evolutions with a proof of the well-posedness and global existence in time and a few illustrative simulations. Last we develop in section 7 a slightly more general picture of shape splines on homogeneous spaces extending the model beyond the landmark setting.

## 2. HAMILTONIAN EQUATIONS OF GEODESICS FOR LANDMARK MATCHING

The problem of landmark matching via diffeomorphic transport is now quite well understood. The basic idea is to build a continuous path of minimal length  $\phi_t$  starting from  $\phi_0 = \text{Id}_{\mathbb{R}^d}$  in a group  $G_V$  of diffeomorphisms, between an initial configuration  $x = (x_i)_{1 \leq i \leq n}$  of  $n$  landmarks in  $\mathbb{R}^d$  and a target configuration  $y = (y_i)_{1 \leq i \leq n}$ . Thus, if  $x_{i,t} \doteq \phi_t(x_i)$  and  $x_t = (x_{i,t})_{1 \leq i \leq n}$ ,  $x_t$  is a path from  $x$  to  $y$  in the space of landmark configurations induced by a geodesic in a Riemannian space of diffeomorphisms. The group  $G_V$  is defined through the flow of time dependent velocity fields  $(t, x) \rightarrow v_t(x)$  on  $\mathbb{R}^d$

$$(1) \quad \frac{\partial \Phi}{\partial t} = v_t \circ \Phi$$

where  $V$  is a Hilbert space of velocity fields and  $v \in L^2([0, 1], V)$  is an element of the space of time dependent velocity fields with finite  $L^2$  norm. Obviously, the existence of a flow  $t \rightarrow \Phi_t^v$  for  $v \in L^2([0, 1], V)$  solution of (1) depends on some regularity assumptions of the instantaneous velocity fields  $v_t$ , mainly the control of the first order derivatives of  $v_t$ . Assuming this, the group  $G_V$  is defined as  $G_V \doteq \{\Phi_1^v \mid v \in L^2([0, 1], V)\}$  and the diffeomorphic matching problem for landmarks is formulated through the following variational problem

$$(2) \quad \begin{cases} \min \int_0^1 |v_t|_V^2 dt, & v \in L^2([0, 1], V) \\ \text{with} \\ \Phi_1^v(x_i) = y_i, & 1 \leq i \leq n. \end{cases}$$

The problem (2) is well posed as soon as  $V$  is continuously embedded in  $C_0^1(\mathbb{R}^d, \mathbb{R}^d)$  the space of  $C^1$  velocities vanishing at  $\infty$  (such a  $V$  is called an admissible space). Namely, we assume the existence of a constant  $C$  such that for any  $v \in V$ , the following inequality is verified

$$(3) \quad |v|_{1,\infty} \leq C|v|_V.$$

Such an admissible space is a Reproducible Kernel Hilbert Space (RKHS) and is equipped with a kernel  $K_V(z, z')$  playing a key role in defining the structure of the solution of the geodesic emerging from (2). Since the reader may not be familiar with RKHS, let us say in a nutshell that for any  $z, z' \in \mathbb{R}^d$  the kernel  $K_V(z, z')$  is a  $d \times d$  matrix in  $\mathcal{M}_d(\mathbb{R})$ , that  $z \rightarrow K_V(z, z')\alpha' \in V$  for any  $z', \alpha' \in \mathbb{R}^d$ , that  $\langle K_V(\cdot, z')\alpha', K_V(\cdot, z'')\alpha'' \rangle_V = \alpha'^T K_V(z', z'')\alpha''$  (the so called reproducible property) and that  $\langle v, K_V(\cdot, z')\alpha' \rangle_V = \langle v(z'), \alpha' \rangle_V$  for any  $v \in V$ . The main fact is that given  $K_V$ , the space  $V$  is completely defined and one may start from the choice of the kernel  $K_V$  itself to define the space  $V$ . A large number of kernels have been proposed most commonly the Gaussian kernel

$$(4) \quad K_V(z, z') = \exp\left(-\frac{|z - z'|^2}{\lambda^2}\right) \text{Id}_{\mathbb{R}^d}.$$

Even more importantly, the kernel appears explicitly in geodesics emerging from (2) as described by the following Theorem. To ensure the existence of a solution in this theorem, we need to assume that the kernel is positive definite:

$$(5) \quad \left| \sum_{i=1}^l K_V(x_i, \cdot) \alpha_i \right|_V^2 = 0 \Rightarrow \forall i \alpha_i = 0.$$

**Theorem 1** (cf [13]). *The solution  $v \in L^2([0, 1], V)$  exists and satisfies*

$$(6) \quad v_t(z) = \sum_{i=1}^n K_V(z, x_{i,t}) p_{i,t}$$

where  $t \rightarrow (x_t, p_t)$  is solution of

$$(7) \quad \begin{cases} \dot{x}_t = \frac{\partial H_0}{\partial p}(x_t, p_t) \\ \dot{p}_t = -\frac{\partial H_0}{\partial x}(x_t, p_t) \end{cases}$$

with  $H_0(p, x) \doteq \frac{1}{2} \sum_{i,j} p_i^T K_V(x_i, x_j) p_j$  and  $x_0 = x$ .

An interesting fact is that the evolution (7) is a Hamiltonian evolution induced by the **Hamiltonian**  $H_0$  on the system of landmarks  $x = (x_1, \dots, x_n)$ . From a mechanical point of view, the landmarks represent the positions of  $n$  particles in  $\mathbb{R}^d$  and each  $p_i$  is the **momentum** attached to the particle  $x_i$ . During the evolution, the particles interact and the equation  $\dot{p}_i = -\frac{\partial H_0}{\partial x_i}(p, x)$  implies that the time derivative of the momentum of the  $i^{th}$  particle is equal to the **internal forces**  $-\frac{\partial H_0}{\partial x_i}(p, x)$  acting on it. This is a simple extension of the usual **Newton equations**.

**Remark 1.** Note that when there is no interaction between the particles i.e.  $K_V(z, z') = \frac{1_{z=z'}}{m} Id_{\mathbb{R}^d}$ , then  $H_0(p, x) = \frac{1}{2m} \sum_{i=1}^n |p_i|^2$  and the system (7) reduces to  $m\dot{x}_i = p_i$  and  $\dot{p}_i = m\ddot{x}_i = 0$  so that the particles evolve independently along straight lines at constant speed. Though this kernel is not admissible, it is not even continuous, this case can be view as a limit case when the characteristic scale  $\lambda \rightarrow 0$  in (4). When the particles interact, the evolutions of the particles are no longer straight lines. The velocity of a given particle is the aggregate of the contributions coming from every particle as described in (6).

This Hamiltonian formulation can be efficiently retrieved by the application of the Pontryagin Maximum Principle (PMP)[1]. The previous assumption in (5) implies that the following control problem is controllable (since for every  $\mathbf{v} = (v_1, \dots, v_n) \in \mathbb{R}^{dn}$  there exists  $u \in V$  such that  $\mathbf{v} = u \cdot x \doteq (u(x_i))_{1 \leq i \leq n}$ ). For any chosen (initial and final) configurations of landmarks, there exists a path between the two landmarks sets. Following an optimal control point of view, we can introduce the Hamiltonian of the system

$$(8) \quad H(p, x, u) = \langle p, \dot{x} \rangle - \frac{1}{2} \langle u, u \rangle_V = \langle p, u \cdot x \rangle - \frac{1}{2} \langle u, u \rangle_V$$

minimizing  $H$  w.r.t. to  $u$  we obtain:

$$(9) \quad u(\cdot) = k(\cdot, x)p,$$

with the compact notation  $k(\cdot, x)p \doteq \sum_{i=1}^n K_V(\cdot, x_i)p_i$ . Then the minimized Hamiltonian can be written as:

$$(10) \quad H_0(p, x) = \frac{1}{2} \sum_{i=1}^n \langle p_j, K_V(x_j, x_i)p_i \rangle = \frac{1}{2} \langle x, k(q, q)x \rangle.$$

Hence we get the Hamiltonian equations of the previous theorem by directly applying the PMP.

It is rather standard to prove that this approach endows the landmark space (set of groups of distinct points) with a Riemannian metric which is induced from that of the group of diffeomorphisms. More generally once a positive definite kernel is given, it gives rise to a Riemannian metric on the space of landmarks. We then use the term kernel metric to designate such a Riemannian metric on the space of landmarks.

The Hamiltonian framework (also viewed from an optimal control viewpoint) can be generalized to shape spaces where the landmark space  $\mathcal{L}$  is replaced with  $L^2(\mathbb{R}/\mathbb{Z}, \mathbb{R}^2)$  for closed curves in the plane [13] or measures. It has been also generalized to discontinuous images [34]. However when deforming embedded objects in  $\mathbb{R}^d$  through the action of smooth vector fields, exact matching usually impossible. The associated control problem is not controllable any more. This is the reason why the Hamiltonian equations are established for the inexact matching problem which includes a penalty term in the minimization. The effect of this penalty term is often to smooth the structure of the momentum used in the Hamiltonian formulation.

### 3. SPLINE INTERPOLATION ON LANDMARK SPACE: A GENERATIVE GROWTH MODEL

**3.1. The standard growth model.** The usual growth model formulation as derived in [27] was initially defined in the case of images. Assume that one has a continuous time sequence of noisy image data  $I_t^D$ ,  $t \in [0, 1]$ , and an image template  $I_0$ . The basic growth estimation problem is to estimate a path  $\phi_t$ ,  $t \in [0, 1]$  in the diffeomorphic group  $G_V$  (a

sequence of deformations) such that  $I_t^D \simeq \phi_t \cdot I_0$  with  $\phi \cdot I = I \circ \phi^{-1}$ . From a Bayesian perspective, this inverse problem is cast as the minimization of the cost

$$(11) \quad J^I(v) = \frac{1}{2} \int_0^1 |v_t|_V^2 dt + \frac{\lambda}{2} \int_0^1 |I_t^D - \phi_t^v \cdot I_0|^2 dt$$

where  $\phi_t^v$  is constrained to be the flow of  $v$  starting at  $\text{Id}$  at time 0 and  $v \in L^2([0, 1], V)$ . Note that the underlying stochastic model for  $v$  is white noise in  $L^2([0, 1], V)$  so that the associated flow  $\phi$  is a **Kunita flow** [23] with almost nowhere differentiable trajectories  $t \rightarrow \phi_t$ .

In the landmark setting addressed in this paper, assume that we have a sequence  $x_t^D$  of data (where  $x_t^D = (x_{t,i}^D)$  is an  $n$ -tuple of  $d$  dimensional landmarks) and a template  $x_0$ . A straightforward adaptation of the above growth model leads to the new cost to be minimized

$$(12) \quad J^x(v) = \frac{1}{2} \int_0^1 |v_t|_V^2 dt + \frac{\lambda}{2} \int_0^1 |x_t^D - \phi_t^v \cdot x_0|^2 dt$$

where  $\phi \cdot x_0 = (\phi(x_{0,i}))$ , and the squared distance  $|x_t^D - \phi_t^v \cdot x_0|^2$  comes from a standard Gaussian white noise model the on individual landmarks. As above the underlying stochastic growth model is a Kunita flow for  $\phi_t$  with irregular trajectories (see Fig. 7).

However, considering that  $x^D$  is continuously observed in time, the minimization problem (12) can be cast into an optimal control problem associated to the reduced Hamiltonian  $H$  given by [24]

$$H(p, x, t) = H_0(p, x) - \frac{\lambda}{2} |x_t^D - x|^2 \text{ where } H_0(p, x) = \frac{1}{2} \sum_{i,j} p_i^T K(x_i, x_j) p_j.$$

The associated optimal trajectories are solutions of the Hamiltonian flow:

$$(13) \quad \begin{cases} \dot{x}_t = \frac{\partial H_0}{\partial p}(p_t, x_t) \\ \dot{p}_t = -\frac{\partial H_0}{\partial x}(p_t, x_t) + \lambda(x_t - x_t^D) \end{cases}$$

The optimal time dependent vector field  $v$  can be reconstructed through the equality

$$v_t(z) = \sum_i K(z, x_{t,i}) p_{t,i}.$$

When  $\lambda \rightarrow 0$ , the optimal evolution converges to a geodesic evolution given by the usual Hamiltonian  $H_0$ . This is not surprising since in (13) the second term of the right hand side vanishes and leads to the minimization of the kinetic energy.

Note that the filtered trajectory  $x$ , obtained from the observation  $x^D$ , is  $C^1$  in time for continuous time observations. It thus has the advantage of correcting the poor prior stochastic growth model given by Kunita flows. However, this is due to the fact that we are assuming *continuous* time observations. In contrast, for discrete observation times  $t_k$ , the optimal solution is piecewise geodesic, again no more  $C^1$ , and offers limited interpolation properties. A better prior on  $x$  is needed.

**3.2. Perturbative growth model.** Denoting  $u_t = \lambda(x_t - x_t^D)$  the perturbation from the geodesic evolution, the estimated trajectory  $x$  extracted from the noisy observation  $x_t^D$ ,  $t \in [0, 1]$  can be reconstructed from the integration of

$$(14) \quad \begin{cases} \dot{x}_t = \frac{\partial H_0}{\partial p}(p_t, x_t) \\ \dot{p}_t = -\frac{\partial H_0}{\partial x}(p_t, x_t) + u_t \end{cases}$$

derived from the ODE (13) if one knows the initial momentum  $p_0$  and  $u$ . The fundamental idea is that this (ODE) can play the role of a generative engine for  $C^1$  trajectories if we put the proper constraint on  $u$  seen as a *control* variable. Interestingly, from a mechanical perspective,  $u_t$  can be interpreted as **external forces** acting on the landmark configuration ( $u_{t,i}$  is the force acting on particle  $i$ ). In absence of external force, the landmark configuration follows a geodesic evolution. When the external forces do not vanishing, the trajectory followed by these landmarks deviates from a simple geodesic to various interpolation trajectories. The question of the dynamics of these external forces is of importance.

The first step is to note that when  $u_t = \lambda(x - x_t^D)$ , the cost of the filtered trajectory  $x$  can be written as

$$(15) \quad J^x(v) = \frac{1}{2} \int_0^1 |v_t|_V^2 dt + \frac{1}{2\lambda} \int_0^1 |u_t|^2 dt,$$

so the minimization of  $J^x$  provides a control on both terms of the right hand side, both positive quantities. Moreover, for a given distribution  $P_{x^D}$  on  $x^D$ , the associated distribution  $P_x$  on the filtered trajectories  $x$  can be defined through the generative model (13) for the adequate distribution  $P_0$  on  $(p_0, u)$ . Very little can be said on the distribution  $P_0$ , but as a first approximation we can proceed as follows. If the overall model minimization of (15) for any observation  $x^D$  provides a reasonable filtered trajectory  $x$  and since for such  $x$ , the quantity  $\frac{1}{2\lambda} \int_0^1 |u_t|^2 dt \leq J^x(v)$  should stay small (at least controlled), one can put a **constraint** on its expectation under  $P_0$

$$(16) \quad E_0 \doteq E_{P_0} \left( \int_0^1 |u_t|^2 dt \right).$$

Using no more information and maximizing entropy [10, 20] under the constraint (16), a reasonable first order model for the marginal distribution of  $u$  is given by a **standard white noise** i.e.  $u_t = \sigma dB_t$  where it is reasonable to set  $\sigma^2$  close<sup>1</sup> to the value  $\lambda$ .

The marginal distribution on  $p_0$  is more problematic, as well as the conditional distribution of  $u$  given  $p_0$ . The simplest solution is to assume independence of  $p_0$  and  $u$ :

$$P_0(dp_0, du) = P_0(dp_0) \otimes P_0(du)$$

so that that conditional generative model ( $p_0$  fixed) for the trajectory  $x$  can be defined as the *SDE*

$$(17) \quad \begin{cases} dx_t = \frac{\partial H_0}{\partial p}(p_t, x_t) dt \\ dp_t = -\frac{\partial H_0}{\partial x}(p_t, x_t) dt + \sigma dB_t \end{cases}$$

**3.3. New growth estimation minimization problem.** This new growth model generates  $C^1$  solutions and can be the foundation for a new formulation of the growth estimation problem in the realistic situation of sparse discrete observation times. Indeed, if we consider  $M$  observation times  $t_1, \dots, t_M$ , this new prior can be used to derive a new growth estimation method :

$$(18) \quad \left| \begin{aligned} \inf_{p_0, u} J(p_0, u) &\doteq \left\{ E(p_0) + \frac{1}{2} \int_0^1 |u_t|^2 dt + \gamma \sum_{k=1}^M |x_{t_k}^D - x_{t_k}|^2 \right\} \\ &\text{subject to } (x, p) \text{ solution of the ODE (14)} \end{aligned} \right|$$

where  $(p_0, u) \in \mathbb{R}^{nd} \times L^2([0, 1], \mathbb{R}^{nd})$  with initial conditions  $(x_0, p_0)$  and  $E$  comes from the log-likelihood of the prior  $P_0(dp_0)$  on  $p_0$ .

**Remark 1.** *The choice of the regularization term  $E$  for  $p_0$  does not seem important in the finite dimensional case we are considering here. An improper flat prior can be used (as we do in the application below) and the term can be dropped. Alternatively consider  $p_0$  fixed to 0. The weight on  $u$  is however of particular importance since it controls the deviation from a geodesic evolution corresponding to  $u \equiv 0$ .*

One can consider more general penalties on  $u$  such as

$$\int_0^1 \langle K_{x_t} u_t, u_t \rangle dt$$

for a state dependent metric ( $K_x$  is assumed here to be a positive symmetric matrix). However, there should be some rationale for the final choice.

At this point we should mention the Riemannian cubic spline point of view developed in [29]. In this framework, we start from a finite dimensional Riemannian manifold  $M$

---

<sup>1</sup>The different approximations done in this derivation should encourage us to be a little cautious about any strong statement!

and the basic problem is to interpolate between two configurations  $(x_0, \dot{x}_0)$  and  $(x_1, \dot{x}_1)$  with a smooth curve  $\gamma$  minimizing an extended *bending* energy :

$$\begin{cases} \min_{\gamma} B(\gamma) \doteq \int_0^1 |\nabla_{\dot{\gamma}} \dot{\gamma}|_{\gamma}^2 dt \\ \text{subject to} \\ \gamma(0) = x_0, \dot{\gamma}(0) = \dot{x}_0, \gamma(1) = x_1 \text{ and } \dot{\gamma}(1) = \dot{x}_1 \end{cases}$$

where  $\nabla$  is the Levi-Civita connexion and  $|\cdot|_{\gamma}$  is the metric given by the manifold at the current location  $\gamma$ . It is quite clear that in the situation  $M = \mathbb{R}^d$  with the flat Euclidean metric, we get  $B(\gamma) = \int_0^1 |\dot{\gamma}|^2 dt$  corresponding to the classical cubic splines [31, 2]. Moreover, one can check (see appendix) that in our Hamiltonian framework, we have

$$(19) \quad u = \dot{p} + \partial_x H_0 = K_x^{-1} \nabla_{\dot{x}} \dot{x}$$

where  $K_x \doteq (K_V(x_i, x_j))_{ij}$  and  $K_x^{-1}$  is the metric tensor at the current landmark positions. Thus, the bending energy is given in our situation as a function of  $u$  :

$$B(u) = \int_0^1 \langle K_x u, u \rangle dt = \int_0^1 |u|_{x,*}^2 dt$$

where  $|u|_{x,*}$  is the induced metric on the cotangent space.

**Remark 2.** *This penalization is different from the simpler  $L^2$  norm we derived previously. On one hand, the Riemannian cubic splines are completely defined by a unique metric underlying the Riemannian structure of the manifold, keeping a pure geometric and intrinsic point of view. However, this intrinsic point of view is not as natural as it may look at first glance since it links tightly the internal forces (given by the term  $\partial_x H_0$ ) and the external forces (given by  $u$ ) to the same metric structure. This is quite questionable since the internal (resp. the external) forces proceed from intrinsic (resp. extrinsic) phenomena.*

We advocated in the previous subsection to *link the choice of the metric on the control  $u$  to the noise model on the observation data*. We derived the particular case of the white noise situation, which produces the standard Euclidean flat metric on the control. However, more general situations could be of some interest leading to non-standard metrics on  $u$ .

**3.4. A non standard metric.** Assume that there exist a Hilbert space  $W$  and a smooth mapping  $\psi : \mathbb{R}^{nd} \rightarrow W^*$  such that the data term in (15) is replaced by

$$\frac{\lambda}{2} \int_0^1 |\psi(x_t^D) - \psi(x_t)|_{W^*}^2 dt,$$

where  $|\cdot|_{W^*}$  is the dual norm on the dual space  $W^*$ . This situation is quite natural if we consider that the true observed quantity is not  $x^D$  but an element  $\psi(x^D) \in W^*$ . This setting is of particular importance in the situation of shape modelling where the individual label of the different landmarks cannot be observed and the point-wise correspondence between the template and the observation is problematic. The so-called measure framework developed in [12] makes intensive use of such a data term where  $\psi(x) = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$  is the empirical distribution of the landmarks and  $W$  is a proper Reproducible Kernel Hilbert Space.

In this new case, the associated optimal trajectories are solutions of the Hamiltonian flow :

$$(20) \quad \begin{cases} \dot{x}_t = \frac{\partial H_0}{\partial p}(p_t, x_t) \\ \dot{p}_t = -\frac{\partial H_0}{\partial x}(p_t, x_t) + \lambda(\psi'(x_t))^{\dagger}(\psi(x_t^D) - \psi(x_t)) \end{cases}$$

where  $\psi'(x_t)^{\dagger} : W^* \rightarrow \mathbb{R}^{nd}$  is the Hilbertian adjoint of the differential  $\psi'(x_t) : \mathbb{R}^{nd} \rightarrow W^*$  of  $\psi$  at  $x$ . As above, denoting  $u_t = \lambda(\psi'(x_t))^{\dagger}(\psi(x_t^D) - \psi(x_t))$  and assuming that  $\psi'(x_t)^{\dagger} \psi'(x_t) : \mathbb{R}^{nd} \rightarrow \mathbb{R}^{nd}$  is invertible (i.e.  $\psi'(x_t)$  is one to one) along the trajectory  $x_t$ ,

we get

$$\begin{aligned} \frac{\lambda}{2} \int_0^1 |\psi(x_t^D) - \psi(x_t)|_{W^*}^2 dt &\geq \frac{\lambda}{2} \int_0^1 |\pi_{x_t}(\psi(x_t^D) - \psi(x_t))|_{W^*}^2 dt \\ &= \frac{1}{2\lambda} \int_0^1 \langle [\psi'(x_t)^\dagger \psi'(x_t)]^{-1} u_t, u_t \rangle dt, \end{aligned}$$

where  $\pi_x$  is the orthogonal projection in  $W^*$  on  $\text{Im}(\psi'(x))$  which is the tangent space at  $\psi(x)$  of the sub-manifold of  $W^*$  defined by the immersion  $\psi$ . In particular, we can use as local metric  $K_x = [\psi'(x)^\dagger \psi'(x)]^{-1}$  which is now position dependent and defines a new Riemannian metric on the Landmark space.

In the mentioned situation of measure representation of landmarks we have the following proposition :

**Proposition 1.** *Assume that the kernel  $K_W$  associated with the RKHS  $W$  is  $C^2$ . Then  $\psi : \mathbb{R}^{nd} \rightarrow W^*$  defined by  $\psi(x) = \frac{1}{n} \sum_{i=1}^n \delta_{x_i}$  is differentiable and  $[\psi'(x)^\dagger \psi'(x)]$  is the block diagonal matrix :*

$$[\psi'(x)^\dagger \psi'(x)] = \frac{1}{n^2} \left( \frac{\partial^2 K_W}{\partial x_i \partial x_j}(x_i, x_j) \right)_{1 \leq i, j \leq n}.$$

*Proof.* See appendix. □

#### 4. EXISTENCE RESULTS AND EULER-LAGRANGE EQUATION FOR SHAPE SPLINE ESTIMATION

In this section we provide several rigorous results concerning the shape spline framework. Indeed, although the overall picture is formally quite clear, it is necessary to provide rigorous statements about the main objects we have just introduced. In particular, conditions for the existence of solutions, global in time, of the perturbed evolution (14) need to be specified, as well as an existence theorem for the growth estimation minimization problem (18). On a more practical side, the minimization problem will be solved by gradient descent, for which we will compute the directional derivatives of functional  $J$  and the associated Euler-Lagrange equations.

**4.1. Existence of controlled evolution and weak dependency in the control variable.** First define the function  $f(q, u) \doteq (\partial_p H_0(x, p), -\partial_x H_0(x, p) + u)^T$  for  $q = (x, p) \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$  and  $u \in \mathbb{R}^{nd}$ . We will consider the following hypothesis on  $V$  and its kernel  $K_V$ :

**H0 :**  *$V$  is continuously embedded in  $C_0^1(\mathbb{R}^d, \mathbb{R}^d)$  and its kernel  $K_V$  is  $C^2$  in each of its variable.*

Following is the first result on the existence in time of a solution to the perturbed evolution (14).

**Proposition 2** (Existence of solutions, global in time, of the controlled system). *Assume (H0). Then for any  $u \in L^2([0, T], \mathbb{R}^{nd})$  and for any initial condition  $q_0 = (x_0, p_0) \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$ , there exists  $q = (x, p) \in C([0, T], \mathbb{R}^{nd} \times \mathbb{R}^{nd})$  such that*

$$(21) \quad \begin{cases} x_t = x_0 + \int_0^t \partial_p H_0(x_s, p_s) ds \\ p_t = p_0 + \int_0^t (-\partial_x H_0(x_s, p_s) + u_s) ds. \end{cases}$$

Moreover, there exists  $C_1, C_2, C_3 > 0$  independent of  $u$  and  $q_0$  such that for any  $t \leq T$  we have

- (1)  $H_0(x_t, p_t) \leq \tilde{H}_T \doteq C_1(H_0(x_0, p_0) + T \int_0^T |u_s|^2 ds),$
- (2)  $|x_t| \leq |x_0| + C_2 T \tilde{H}_T^{1/2},$
- (3)  $|p_t| \leq (|p_0| + \int_0^T |u_s| ds) \exp(C_3 T \tilde{H}_T^{1/2}).$

*Proof.* Under (H0), we have existence and uniqueness locally in time of the solution to the system (21) for any initial condition  $q_0 = (x_0, p_0)$ . The only point to be checked is that the solution does not go to infinity in finite time. Let  $t_0 < T$  be such that we have a solution  $q_t = (x_t, p_t)$  defined on  $[0, t_0[$ . From (21), we get that  $H_t \doteq H_0(x_t, p_t) = H_0(x_0, p_0) + \int_0^t \langle \partial_p H_0(x_s, p_s), u_s \rangle ds$ . From (3), since  $\partial_{p_i} H_0(x_s, p_s) = v(x_i)$  for  $v(\cdot) = \sum_{j=1}^n K_V(\cdot, x_j) p_j$ ,

we get  $|\partial_p H_0(x_s, p_s)| \leq C H_0(x_s, p_s)^{1/2}$  for a universal constant  $C$  and we deduce that  $H_t \leq H_0(x_0, p_0) + C(\max_{s \in [0, t]} H_s)^{1/2} \int_0^t |u_s| ds$ . Hence

$$(22) \quad \max_{s \in [0, t]} H_s \leq 2H_0(x_0, p_0) + 4C^2 \left( \int_0^t |u_s| ds \right)^2 \leq 2H_0(x_0, p_0) + 4C^2 T \int_0^T |u_s|^2 ds \leq \tilde{H}_T$$

for  $C_1 = \max(2, 4C^2)$ . The upper bound does not depend on  $t_0$  and in particular the Hamiltonian can not explode in finite time. It is sufficient now to prove that  $x_t$  and  $p_t$  stay also bounded. Indeed, we have

$$(23) \quad |x_t| \leq |x_0| + \left| \int_0^t \partial_p H_0(x_s, p_s) ds \right| \leq |x_0| + C \int_0^T H_s^{1/2} ds \leq |x_0| + CT \tilde{H}_T^{1/2}.$$

Moreover,  $|p_t| \leq \int_0^t |\partial_x H_0(x_s, p_s)|_\infty |p_s| ds + \int_0^t |u_s| ds + |p_0|$  and since again  $|\partial_{x_i} H_0(x, p)| = |dv^*(x_i) p_i|$  for  $v(\cdot) = \sum_{j=1}^n K_V(\cdot, x_j) p_j$ , we get from (3) that there exists  $C' > 0$  such that  $|p_t| \leq C' \int_0^t H_s^{1/2} |p_s| ds + \int_0^t |u_s| ds + |p_0|$ . Using Gronwall's Lemma, we get eventually

$$(24) \quad |p_t| \leq (|p_0| + \int_0^T |u_s| ds) \exp(C' \int_0^T H_s^{1/2} ds) \leq (|p_0| + \int_0^T |u_s| ds) \exp(C' T \tilde{H}_T^{1/2}).$$

From (22), (23) and (24) we deduce that  $x_t$  and  $p_t$  stay uniformly bounded on  $[0, t_0]$ . Since  $t_0 < T$  is arbitrary, the system (21) admits a solution on  $[0, T]$  satisfying points 1), 2) and 3) of Proposition 2.  $\square$

**Proposition 3** (Dependence in  $u$ ). *Assume (H0). For any  $q_0 \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$  and any  $u \in L^2([0, T], \mathbb{R}^{nd})$ , let  $q^{u, q_0} \in C([0, T], \mathbb{R}^{nd} \times \mathbb{R}^{nd})$  denote the solution of (21) with initial condition  $q_0$ . Then the mapping  $(q_0, u) \rightarrow q^{u, q_0}$  is continuous for the weak topology on  $L^2([0, T], \mathbb{R}^{nd})$  and uniform convergence on  $C([0, T], \mathbb{R}^{nd} \times \mathbb{R}^{nd})$ .*

*Proof.* Let  $u^n \rightharpoonup u^\infty$  be a weakly converging sequence in  $L^2$  and  $q_0^n \rightarrow q_0^\infty$  be a converging sequence of initial conditions. There exists  $R > 0$  such that  $\sup_{n \geq 0} |u^n|_2 \leq R$  and by Proposition 2, if  $q^n$  (resp.  $q^\infty$ ) denotes the solution of (21) for  $(q_0, u) = (q_0^n, u^n)$  (resp.  $(q_0, u) = (q_0^\infty, u^\infty)$ ), then there exists  $M > 0$  such that  $|q_t^n| \leq M$  for  $(t, n) \in [0, T] \times \mathbb{N} \cup \{\infty\}$ . Let  $K_M > 0$  such that

$$|f(q, u) - f(q', u')| \leq K_M |q - q'| + |u - u'|$$

for  $|q|, |q'| \leq M$  and  $u, u' \in \mathbb{R}^{nd}$  (such  $K_M$  exists since (H0) implies that  $dH_0$  is  $C^1$ ). We have

$$\begin{aligned} |q_t^n - q_t^\infty| &= \left| \int_0^t f(q_s^n, u_s^n) - f(q_s^\infty, u_s^\infty) ds + (q_0^n - q_0^\infty) \right| \\ &\leq K_M \int_0^t |q_s^n - q_s^\infty| ds + \left| \int_0^t (u_s^n - u_s^\infty) ds \right| + |q_0^n - q_0^\infty| \end{aligned}$$

so that using Gronwall's Lemma we get

$$(25) \quad |q_t^n - q_t^\infty| \leq (|q_0^n - q_0^\infty| + \sup_{r \leq T} \left| \int_0^r (u_s^n - u_s^\infty) ds \right|) \exp(K_M T).$$

Since  $u \rightarrow \int_0^r u_s ds$  is a continuous linear mapping, we get from the weak convergence that  $|\int_0^r (u_s^n - u_s^\infty) ds| \rightarrow 0$  as  $n \rightarrow \infty$ . Noticing that  $|\int_r^{r'} (u_s^n - u_s^\infty) ds| \leq 2\sqrt{r' - r}R$  for any  $r < r' \leq T$ , Ascoli's Theorem turns the previous simple convergence into uniform convergence in  $r$ .  $\square$

**Theorem 2** (Existence of a minimizer in the inexact case). *Assume (H0) and that  $q \rightarrow \mathcal{K}_q \in \mathcal{M}_d(\mathbb{R})$  is  $C^0$  with  $\langle \mathcal{K}_q u, u \rangle \geq c|u|^2$  for some fixed  $c > 0$  and that  $C(q, u) = \frac{1}{2} \langle \mathcal{K}_q u, u \rangle$ . Assume that  $q_0 \rightarrow E(q_0)$  is a non negative and lower semi-continuous function such that  $E \rightarrow +\infty$  when  $|q_0| \rightarrow \infty$ , that  $g_k$  is a continuous non negative function for any  $1 \leq k \leq M$  and let  $x_0 \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$  be a fixed initial condition. Then the function*

$$J(q_0, u) \doteq E(q_0) + \int_0^T C(q_t^{q_0, u}, u_t) dt + \sum_{k=1}^M g_k(q_{t_k}^{q_0, u})$$

*defined for any  $(q_0, u) \in \mathbb{R}^{nd} \times L^2([0, T], \mathbb{R}^{nd})$  (where  $q^{q_0, u}$  is the solution of (21) with initial condition  $q_0$ ) reaches its minimum.*

*Proof.* Let  $(q_0^n, u^n)_{n \geq 0}$  be a minimizing sequence for  $J$ . Since the  $g_k$ 's are non negative and  $E \geq 0$  with  $E \rightarrow +\infty$  when  $|q_0| \rightarrow +\infty$ , we get that  $\int_0^T \langle \mathcal{K}_{q_s^n} u_s^n, u_s^n \rangle ds$  and  $|q_0^n|$  are upper bounded. Since we assume that  $\langle \mathcal{K}_q u, u \rangle \geq c|u|^2$ , we deduce that  $u^n$  is a bounded sequence in  $L^2$  and we get by weak compactness of the strong balls that (up to the extraction of a sub-sequence) there exists  $u^\infty$  such that  $u^n \rightharpoonup u^\infty$ . Again, up to the extraction of a subsequence, we can assume that there exists  $q_0^\infty$  such that  $q_0^n \rightarrow q_0^\infty$ . Since

$$\begin{aligned} \int_0^T \langle \mathcal{K}_{q_s^n} u_s^n, u_s^n \rangle ds &\geq \int_0^T \langle \mathcal{K}_{q_s^n} - \mathcal{K}_{q_s^\infty} \rangle u_s^n, u_s^n \rangle ds + \int_0^T \langle \mathcal{K}_{q_s^\infty} u_s^\infty, u_s^\infty \rangle ds \\ &+ 2 \int_0^T \langle \mathcal{K}_{q_s^\infty} u_s^\infty, u_s^n - u_s^\infty \rangle ds, \end{aligned}$$

we deduce from Proposition 3 that

$$\sup_{s \leq T} \|\mathcal{K}_{q_s^n} - \mathcal{K}_{q_s^\infty}\| \rightarrow 0 \text{ and } \int_0^T \langle (\mathcal{K}_{q_s^n} - \mathcal{K}_{q_s^\infty}) u_s^n, u_s^n \rangle ds \rightarrow 0.$$

Moreover, by weak convergence  $\int_0^T \langle \mathcal{K}_{q_s^\infty} u_s^\infty, u_s^n - u_s^\infty \rangle ds \rightarrow 0$  so that  $\int_0^T C(q_s^\infty, u_s^\infty) ds \leq \liminf \int_0^T C(q_s^n, u_s^n) ds$  and by continuity of the  $g_k$ 's and of  $E$ ,  $J(q_0^\infty, u^\infty) \leq \lim J(q_0^n, u^n) = \inf J$ .  $\square$

Obviously, this proof gives also the existence of a minimizer in the *exact* case where the spline is constrained to go through a sequence  $(x_{t_k}^D)_{1 \leq k \leq M}$  of landmarks configurations if there exists at least one such controlled path with finite cost.

#### 4.2. Directional derivatives and Euler-Lagrange Equation.

**Proposition 4.** Assume (H0) and let  $u, \delta u \in L^2([0, T], \mathbb{R}^{nd})$  and  $q_0, \delta q_0 \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$ . For any  $\epsilon > 0$ , we denote  $q_{t,\epsilon} \in C([0, T], \mathbb{R}^{nd} \times \mathbb{R}^{nd})$ , the solution of (21) for the control  $u_{t,\epsilon} \doteq u_t + \epsilon \delta u_t$  and the initial condition  $q_{0,\epsilon} = q_0 + \epsilon \delta q_0$ . Then we have

$$(26) \quad \limsup_{\substack{\epsilon \neq 0 \\ \epsilon \rightarrow 0}} \sup_{t \leq T} \left| \frac{q_{t,\epsilon} - q_t}{\epsilon} - \delta q_t \right| = 0$$

where  $\delta q$  is the absolutely continuous solution on  $[0, T]$  of the linearized system

$$(27) \quad \dot{\delta q}_t = \partial_q f(q_t, u_t) \delta q_t + \partial_u f(q_t, u_t) \delta u_t$$

with initial condition  $\delta q_0$ .

*Proof.* Assume  $0 < \epsilon \leq 1$  so that  $\sup_{0 < \epsilon \leq 1} \int_0^T |u_{s,\epsilon}|^2 dt \leq 2 \int_0^T |u_s|^2 + |\delta u_s|^2 ds < \infty$ . As we did in the proof of Proposition 3 we deduce from Proposition 2 that  $q_{t,\epsilon} \doteq (x_{t,\epsilon}, p_{t,\epsilon})$  is uniformly bounded by some  $M > 0$  for  $(t, \epsilon) \in [0, T] \times ]0, 1]$ . Again, if  $K_M > 0$  is such that

$$|f(q, u) - f(q', u')| \leq K_M |q - q'| + |u - u'|$$

for  $|q|, |q'| \leq M$  and  $u, u' \in \mathbb{R}^{nd}$  then  $|q_{t,\epsilon} - q_t| \leq K_M \int_0^t |q_{s,\epsilon} - q_s| ds + \epsilon (\delta q_0 + \int_0^t |\delta u_s| ds)$  and by Gronwall's Lemma

$$(28) \quad |q_{t,\epsilon} - q_t| \leq \epsilon \exp(K_M T) (\delta q_0 + \int_0^T |\delta u_s| ds).$$

Since from (H0),  $f$  is  $C^1$  and  $\partial_q f(q, u)$  and  $\partial_u f(q, u)$  are uniformly bounded for  $|q| \leq M$ , there exists a unique solution  $\delta q$  of (27) which is absolutely continuous. Moreover,

$$(29) \quad \begin{aligned} &\left| \frac{q_{t,\epsilon} - q_t}{\epsilon} - \delta q_t \right| \\ &\leq \int_0^t \left| \frac{f(q_{s,\epsilon}, u_{s,\epsilon}) - f(q_s, u_s)}{\epsilon} - \partial_q f(q_s, u_s) \delta q_s + \partial_u f(q_s, u_s) \delta u_s \right| ds \end{aligned}$$

$$(30) \quad \leq \int_0^t \left| \partial_q f(q_s, u_s) \left( \frac{q_{s,\epsilon} - q_s}{\epsilon} - \delta q_s \right) \right| ds + \int_0^t |\eta_{s,\epsilon}| ds$$

where

$$\begin{aligned}\eta_{s,\epsilon} &\doteq \left| \frac{f(q_{s,\epsilon}, u_{s,\epsilon}) - f(q_s, u_s)}{\epsilon} - \partial_q f(q_s, u_s) \frac{q_{s,\epsilon} - q_s}{\epsilon} + \partial_u f(q_s, u_s) \delta u_s \right| \\ &= \left| \frac{f(q_{s,\epsilon}, u_s) - f(q_s, u_s)}{\epsilon} - \partial_q f(q_s, u_s) \frac{q_{s,\epsilon} - q_s}{\epsilon} \right|.\end{aligned}$$

However, from (28) and the fact that  $\partial_q f$  is uniformly bounded for  $|q| \leq M$ , we get that  $\eta_{s,\epsilon}$  is uniformly bounded on  $[0, T] \times ]0, 1]$ . Since  $\eta_{s,\epsilon} \rightarrow 0$  for  $\epsilon \rightarrow 0$ , we get by Lebesgue's Dominated Convergence Theorem that  $\int_0^T \eta_{s,\epsilon} ds \rightarrow 0$ . Using (29) and Gronwall's Lemma, we get

$$\left| \frac{q_{t,\epsilon} - q_t}{\epsilon} - \delta q_t \right| \leq \left( \int_0^T \eta_{s,\epsilon} ds \right) \exp\left( \int_0^T \partial_q f(q_s, u_s) ds \right) \rightarrow 0.$$

□

**Theorem 3** (Directional derivative). *Assume (H0), assume that  $q \rightarrow \mathcal{K}_q \in \mathcal{M}_d(\mathbb{R})$  is  $C^1$ , that  $g_k$  is  $C^1$  for any  $1 \leq k \leq M$  and that  $E$  is  $C^1$ . Let  $C(q, u) = \frac{1}{2} \langle \mathcal{K}_q u, u \rangle$  and  $T > t_M$ . Then for any  $u, \delta u \in L^2([0, T], \mathbb{R}^{nd})$  and  $q_0, \delta q_0 \in \mathbb{R}^{nd} \times \mathbb{R}^{nd}$ , if*

$$J(\epsilon) \doteq E(q_{0,\epsilon}) + \int_0^T C(q_{t,\epsilon}, u_{t,\epsilon}) dt + \sum_{k=1}^M g_k(q_{t_k,\epsilon})$$

we have

$$(31) \quad \lim_{\epsilon \rightarrow 0} \frac{J(\epsilon) - J(0)}{\epsilon} = \langle \nabla E, \delta q_0 \rangle + \int_0^T (\partial_q C(q_s, u_s) \delta q_s + \partial_u C(q_s, u_s) \delta u_s) ds + \sum_{k=1}^M \langle \nabla g_k(q_{t_k}), \delta q_{t_k} \rangle$$

$$(32) \quad = \langle \nabla E(q_0) + P_0, \delta q_0 \rangle + \int_0^T \langle \nabla_u C(q_s, u_s) + \partial_u f^T(q_s, u_s) P_s, \delta u_s \rangle ds$$

where  $C(q, u) \doteq \langle \mathcal{K}_q u, u \rangle$ ,  $\delta q_t$  is solution of (27) and  $P_t$  is of bounded variations with  $P_T = 0$  and

$$(33) \quad dP_t = -\partial_q f(q_t, u_t)^T P_t dt - \sum_{k=1}^M \nabla g_k(q_{t_k}) \otimes \delta_{t_k}$$

where  $v \otimes \delta_x$  denotes a vectorial Dirac measure at location  $x$  with value  $v$ .

*Proof.* For any variation  $\delta u$  of the control  $u$ , we get by Proposition 4 that  $\epsilon \rightarrow q_{t,\epsilon}$  is differentiable and  $\partial_\epsilon q|_{\epsilon=0} = \delta q$  where  $\delta q = \partial_q f(q_s, u_s) \delta q_s + \partial_u f(q_s, u_s) \delta u_s$ . Moreover,

$$\begin{aligned}A(\epsilon) &\doteq \frac{J(\epsilon) - J(0)}{\epsilon} \\ &= \left( \langle \nabla E(q_0), \delta q_0 \rangle + \int_0^T (\partial_q C(q_s, u_s) \delta q_s + \partial_u C(q_s, u_s) \delta u_s) ds + \sum_{k=1}^M \langle \nabla g_k(q_{t_k}), \delta q_{t_k} \rangle \right) \\ &= \int_0^T \partial_q C(q_s, u_s) \left( \frac{q_{s,\epsilon} - q_s}{\epsilon} - \delta q_s \right) ds \\ &\quad + \sum_{k=1}^M \langle \nabla g_k(q_{t_k}), \frac{q_{t_k,\epsilon} - q_{t_k}}{\epsilon} - \delta q_{t_k} \rangle + \zeta_\epsilon + \int_0^T \eta_{s,\epsilon} ds\end{aligned}$$

where

$$\zeta_\epsilon = \sum_{k=1}^M \left( \frac{g_k(q_{t_k,\epsilon}) - g_k(q_{t_k})}{\epsilon} - \langle \nabla g_k(q_{t_k}), \frac{q_{t_k,\epsilon} - q_{t_k}}{\epsilon} \rangle \right) + \frac{E(q_{0,\epsilon}) - E(q_0)}{\epsilon} - \langle \nabla E(q_0), \delta q_0 \rangle$$

and

$$\eta_{s,\epsilon} = \frac{C(q_{s,\epsilon}, u_{s,\epsilon}) - C(q_s, u_s)}{\epsilon} - \left( \partial_q C(q_s, u_s) \frac{q_{s,\epsilon} - q_s}{\epsilon} + \partial_u C(q_s, u_s) \delta u_s \right).$$

From the fact that the  $g_k$ 's and  $E$  are  $C^1$  and (28), we get that  $\zeta_\epsilon \rightarrow 0$  with  $\epsilon \rightarrow 0$ . Since  $q_{s,\epsilon}$  is uniformly bounded for  $(s, \epsilon) \in [0, T] \times ]0, 1]$  and  $\mathcal{K}_q$  is  $C^1$ , one easily gets from (28)

that there exist  $a, b > 0$  such that  $|\eta_{s,\epsilon}| \leq a + b(|u|_s^2 + |\delta u_s|^2)$ . Since  $\eta_{s,\epsilon} \rightarrow 0$  as  $\epsilon \rightarrow 0$ , we get by Lebesgue's Dominated Convergence Theorem that  $\int_0^T \eta_{s,\epsilon} ds \rightarrow 0$ . Using (26), we get eventually  $A(\epsilon) \rightarrow 0$  so that

$$(34) \quad \begin{aligned} \delta J &\doteq \frac{dJ}{d\epsilon}(0) = \langle \nabla E(q_0), \delta q_0 \rangle \\ &\quad + \int_0^T (\partial_q C(q_s, u_s) \delta q_s + \partial_u C(q_s, u_s) \delta u_s) ds + \sum_{k=1}^M \langle \nabla g_k(q_{t_k}), \delta q_{t_k} \rangle \end{aligned}$$

and (31) is proved. Introducing now  $M_{t,s}$  the semi-group solution of  $\partial_s M_{t,s} = \partial_q f(q_s, u_s) M_{t,s}$  with  $M_{t,t} = \text{Id}_{2nd}$  we get

$$\delta q_s = \int_0^s M_{t,s} \partial_u f(q_t, u_t) \delta u_t dt + M_{0,s} \delta q_0.$$

and from (34)

$$(35) \quad \begin{aligned} \delta J &= \langle \nabla E(q_0), \delta q_0 \rangle + \int_0^T \langle \nabla_u C(q_t, u_t), \delta u_t \rangle dt \\ &\quad + \int_0^T \langle \nabla_q C(q_s, u_s), \delta q_s \rangle ds + \sum_{k=1}^n \langle \nabla g_k(q_{t_k}), \delta q_{t_k} \rangle. \end{aligned}$$

Thus we have

$$\begin{aligned} \delta J &= \int_0^T \langle \nabla_u C(q_t, u_t) + \partial_u f(q_t, u_t)^T \int_t^T M_{t,s}^T \nabla_q C(q_s, u_s) ds \\ &\quad + \sum_{k=1}^n \partial_u f(q_t, u_t)^T M_{t,t_k}^T \nabla g_k(q_{t_k}) \mathbf{1}_{t \leq t_k}, \delta u_t \rangle dt \\ &\quad + \langle \nabla E(q_0) + \int_0^T M_{0,s}^T \nabla_q C(q_s, u_s) ds + \sum_{k=1}^n M_{0,t_k}^T \nabla g_k(q_{t_k}) \mathbf{1}_{t \leq t_k}, \delta q_0 \rangle \\ &= \int_0^T \langle \nabla_u C(q_t, u_t) + \partial_u f(q_t, u_t)^T P_t, \delta u_t \rangle dt + \langle \nabla E(q_0) + P_0, \delta q_0 \rangle. \end{aligned}$$

where

$$P_t \doteq \int_t^T M_{t,s}^T \nabla_q C(q_s, u_s) ds + \sum_{k=1}^n M_{t,t_k}^T \nabla g_k(q_{t_k}) \mathbf{1}_{t \leq t_k}.$$

One easily sees that  $P_t$  is absolutely continuous between the observation times with jumps at the observation times. Moreover, differentiating  $t \rightarrow M_{t,t'} M_{t',t} = \text{Id}$ , we get  $\partial_t M_{t,t'} = -M_{t,t'} \partial_q f(q_t, u_t)$  so that we get (33).  $\square$

From Thm 3, we get immediately the Euler-Lagrange equations for shape splines evolution :

$$(36) \quad \begin{cases} \dot{q}_t = f(q_t, u_t) \\ \mathcal{K}_{q_t} u_t + P_t^p = 0 \\ dP_t = -\partial_q f(q_t, u_t)^T P_t dt - \sum_{k=1}^M \nabla g_k(q_{t_k}) \otimes \delta_{t_k} \end{cases}$$

where  $P_t = (P_t^x \ P_t^p)^T$  with boundary conditions

$$(37) \quad P_T = 0, \nabla E(q_0) + P_0 = 0.$$

**Remark 1.** Note that in our experiments, we will consider that  $x_0$  is fixed and let  $p_0$  be free (sometimes called natural spline in the classical cubic splines framework). This corresponds to a flat prior on  $p_0$  and gives the new boundary conditions

$$(38) \quad P_T = 0, P_0^p = 0.$$

Moreover, the cost functions  $g_k$  will usually not depend on  $p$  so that  $\partial_p g_k(q) = 0$  and  $P_t^p$  is absolutely continuous in time. In particular, we get in this case that for natural splines,  $u_t$  is continuous and vanishes at the boundary of the interval  $[0, T]$ .

## 5. NUMERICAL EXPERIMENTS

In this section we provide preliminary experiments illustrating the behavior of shape splines in simple 2D synthetic experiments. We start with a small family of observation

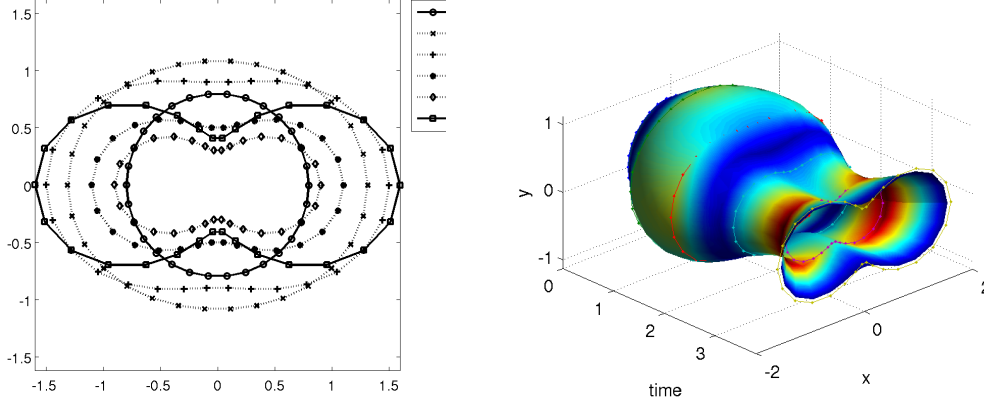


FIGURE 1. Left hand side :2D plots of the sequence of sampled curves  $c_k$  for  $k = 0, \dots, 5$ . The initial curve is a circle and the final shape is a horizontally pinched ellipse (plain curve). The intermediate curves are generated by linear interpolation between the initial and final shapes. Note that no noise is added here and that the intermediate shapes are rescaled by a factor  $r_k$  depending on  $k$ . Right hand side : estimated shape spline displayed in space-time representation. The color indicates the value of the norm of  $u$  on the surface through time (blue for low values, red for high values)

times  $0 \leq t_1 < \dots < t_M \leq T$  and for each time, a landmark configuration  $x_{t_k}^D \in \mathbb{R}^{2n}$  defined as a noisy version of regularly sampled curves  $c_k$ . To stay close to some realistic framework where the observation points are sparse, the number  $M$  of time points is kept small to  $M = 5$  (with  $t_5 = 3.8$ ). In these experiments, we focus on the simple Euclidean metric  $|x_{t_k}^D - x_{t_k}|^2$  in the data term, which is in agreement with our Gaussian noise assumption.

In our first experiment, we consider the evolution of an initial 2D circular shape which somewhat linearly evolves into a pinched ellipsoidal shape. The shapes are regularly sampled in a consistent way so that the problem of point correspondence between two time points does not need to be considered (see Fig 1). To emphasize the space-time regularity provided by the spline shapes interpolation, we display the result as a surface in the 3D space-time. The first interesting point to note is that the shape spline provides an actual smooth interpolation of the evolution in time between the observation epochs and extends to shape spaces the specific behavior of classical cubic splines. In particular, the shape spline actually joins the observed shapes with very good accuracy despite the fact that we are using here inexact matching. A second important point is that the shape spline comes with the estimation in time of the control variable  $u$  which can be interpreted as an external force bending the underlying geodesic. What we see in this example and in the other similar situations, displayed in Fig 2, is that the point-wise values of  $u$  give interesting information on the evolution process. More specifically, in every example, the initial and end shape have specific features (numbers of lobes, orientation, etc) and the most active zones are easily interpretable and correspond to transition regions in the shape evolution.

**5.1. Robustness to noise.** The robustness to noise is a rather important subject from a practical point of view. Indeed, the noise basically degrades the *spatial* resolution

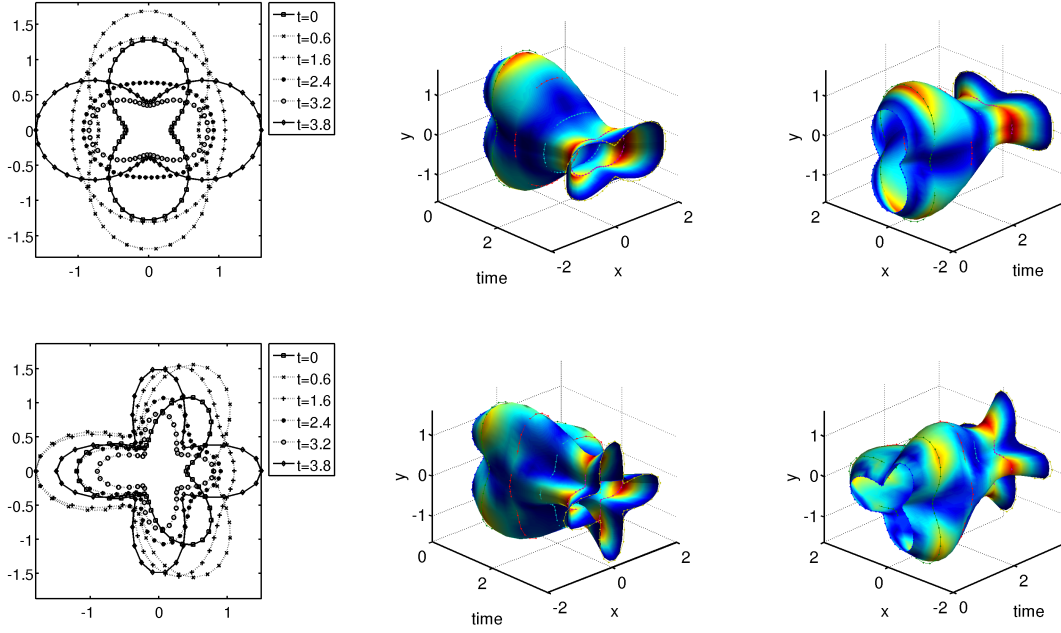


FIGURE 2. The figure displays two estimation experiments of shape splines for two different problems. From left to right, the first column displays the sequence of observed shapes, the second and the third column two different viewpoints. The color of any point on the space-time surface is related to the norm of its control variable (with increasing values from deep blue to red)

of the measurements so that the evolution through time of a particular point of the evolving curve may be a sharply broken line. The standard spline approach can be quite efficient in filtering this noise if the time sampling frequency is high enough. This is hardly the case in many important situations. However, neighboring points behave coherently through time and offer an interesting source of spatial redundancy. Much of the large

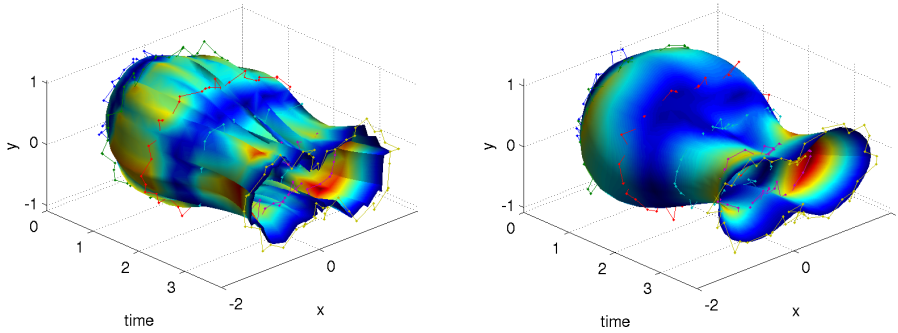


FIGURE 3. Robustness to noise. An i.i.d. Gaussian noise with standard deviation  $\sigma = 0.1$  is added to each measurement point. Shape splines are computed with two different scale parameters,  $\lambda = 0.001$  on the left,  $\lambda = 0.6$  on the right.

deformation shape space theory involves the integration of spatial redundancy in the

comparison between shapes. The shape spline setting, considering shapes as a whole and not as a bag of independent points, keeps this important aspect but adds a time component and considers the problem in the full space-time setting. This robustness to noise is illustrated in Figure 3 where an i.i.d. Gaussian noise with standard deviation  $\sigma = 0.1$  is added and a series of shape splines are computed under increasing values of the spatial regularity scale parameter  $\lambda$  as introduced in (4). For low values of this parameters (with respect to the overall scale of the shapes) the reconstructed evolution is clearly far from any reasonable solution since the spatial redundancy is hardly taken into account. Increasing the value of  $\lambda$  to values in accordance to the scale of the object produces a much better reconstruction of the actual shapes at any observation time but also keeps existing time regularity.

**5.2. Extrapolation.** Another distinguished feature of the usual spline setting which is extended in the shape spline setting is the fact that the extrapolation of the data outside the interval of observation is quite straightforward. Indeed, outside the limits of the

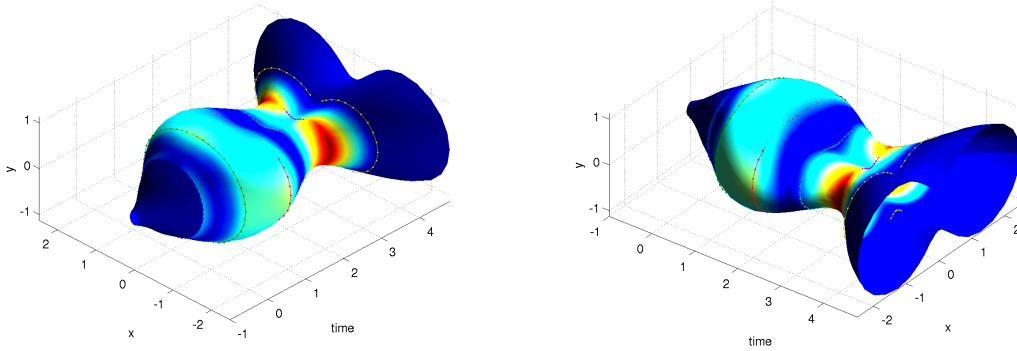


FIGURE 4. Extrapolations. The extrapolation of the evolution at both ends of the observation interval ( $\lambda = 0.6$ ,  $\sigma = 0$ ).

observation interval, the value of the control parameter  $u$  is set to zero and the evolution is naturally extended with a geodesic evolution. Moreover, one can check (see Remark 1) that  $u$  vanishes at the last observation time so that the previous extension is  $C^0$  for the control variable  $u$  and  $C^1$  for the shape variable  $x$ . Note that in the standard growth model described in (12), the evolution is extrapolated by fixing the shape variable to its value at the last observation and this extrapolation is only  $C^0$ . In Figure 4, we display a simple example of extrapolation where the underlying evolution within the observation interval is the same than in Figure 1. The computed extrapolation appears visually quite natural at both ends.

**5.3. Comparison with piecewise geodesic evolution.** We end this section with a comparison with the piecewise geodesic interpolation scheme [27], derived from (12) for a finite set of observation points as the minimizing solution of

$$(39) \quad J^x(v) = \frac{1}{2} \int_0^T |v_t|_V^2 dt + \gamma \sum_{i=1}^M |x_{t_k}^D - x_{t_k}|^2$$

where as previously stated,  $x_t = \phi_t^v(x_0)$  and  $\phi_t^v$  is the flow of  $v \in L^2([0, 1], V)$ . We display an experiment on a synthetic evolution where  $t \rightarrow x_t^D$  is given by a simple analytic formula where a circle evolves smoothly into an ellipse by increasing its eccentricity through time combined with a rotation of the principal axis (see three orthogonal views of the synthetic object in the middle column of Fig. 5). We display the estimated evolution in the piecewise geodesic setting given by (39) and with a shape spline in Fig 5. As expected, the piecewise geodesic estimation provides good results but with a loss of regularity at the observation points as in the simpler situation of piecewise linear interpolation in signal processing. The shape spline seems to perform better at the observation points but also to provide a better estimation *between* observation points. In Fig. 6, we provide a more quantified

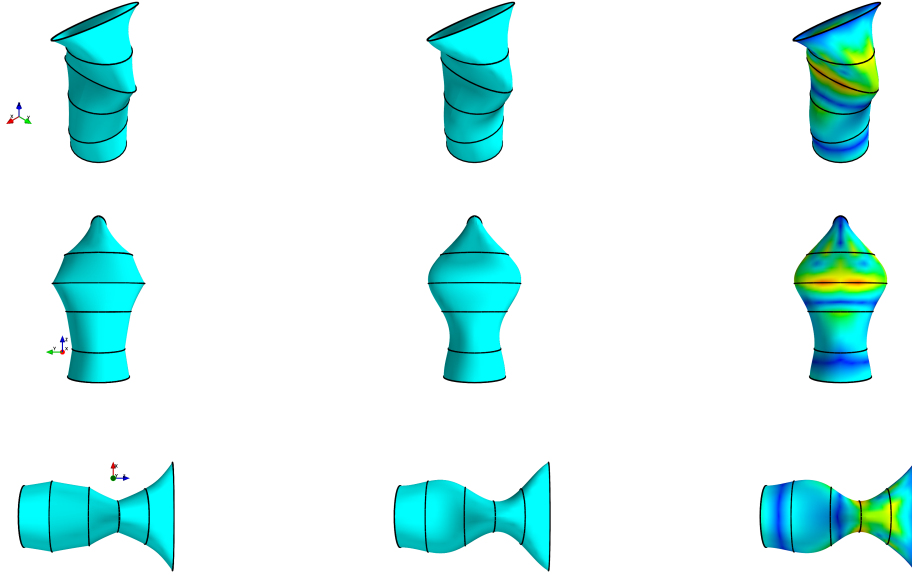


FIGURE 5. Comparison with piecewise geodesic evolution. On this example the middle column corresponds to 3 different views of a synthetic shape evolution (here the time axis is vertical in the first two rows and horizontal in the last row). The first column correspond to the piecewise geodesic evolution one can get from (12) and the last one to the shape spline estimation.

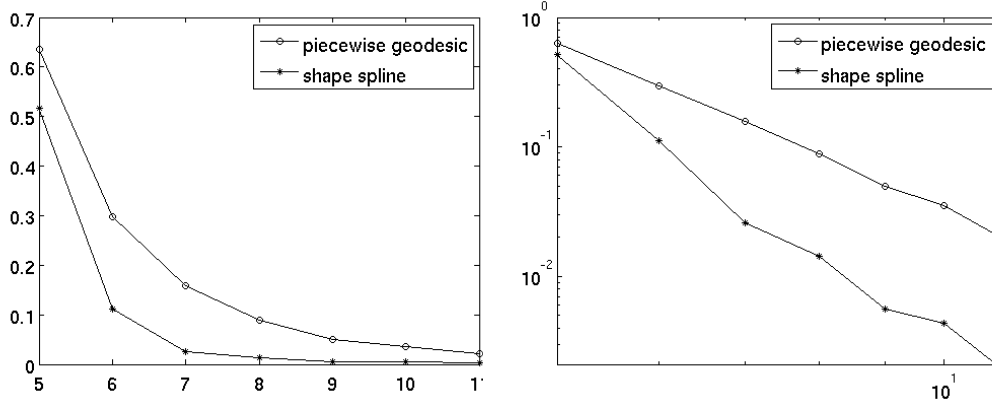


FIGURE 6. Comparison of the  $L^2$  error  $E$  (see (40)) between piecewise geodesic and shape spline estimation. The horizontal axis is the number  $M$  of observation points and the vertical axis the  $L^2$  distance in arbitrary units through time between the estimated shape and the actual shape provided by the synthetic evolution. Right panel, log-log plot.

comparison of the approximation quality between the estimation process by computing the  $L^2$  error

$$(40) \quad E \doteq \left( \int_{t_1}^{t_M} |x_t^D - x_t|^2 dt \right)^{1/2}$$

as a function of the number  $M$  of observations. It is quite clear that the convergence is faster with shape spline interpolation than with the piecewise geodesic interpolation (we do not go beyond 11 observation points since the error is small enough to be approaching other numerical errors in the optimization scheme). The loglog-plot in Fig. 6 seems

to indicate a polynomial convergence in  $C/M^\alpha$  very similar to the classical situation in interpolation theory that ( $\alpha = 2$  for linear spline and  $\alpha = 4$  for cubic spline)<sup>2</sup>

## 6. A STOCHASTIC SHAPE SPLINE MODEL

In this section, we study further the first candidate for the second order model (17) of stochastic evolutions of shapes. We prove that the solutions are well defined for all times and present some simulations that highlight some features of this model.

**6.1. Non blow-up result.** This section is devoted to study the well-posedness of the SDE (17) introduced as our generative growth model in subsection 3.2. The random force is chosen to be an increment of the Brownian motion though we could also have introduced a Levy process in the evolution of the momentum to account for sudden activations of cells. It would turn our growth model into a more realistic one. However, this Brownian perturbation is the first step toward such a model and we will now discuss its feasibility from the mathematical point of view. We will prove that the solutions of the SDE do not blow up in finite time a.s.

The stochastic differential system is:

$$(41a) \quad dp_t = -\partial_x H_0(p_t, x_t) dt + \varepsilon dB_t,$$

$$(41b) \quad dx_t = \partial_p H_0(p_t, x_t) dt.$$

Here,  $\varepsilon$  is a constant parameter and  $B_t$  is a Brownian motion on  $\mathbb{R}^{dn}$ . We will work with the Gaussian kernel but this can be directly extended to other kernels. When  $\varepsilon$  is constant, there is no difference studying these stochastic differential equations with the Ito integral or Stratonovich one. However, for a general variance term, we will use the Ito stochastic integral. From the theorem of existence and uniqueness of solution of stochastic differential equation under the linear growth conditions, we can work on the solutions of such equations for a large range of kernels. Yet in our case the Hamiltonian is quadratic, and the classical results for existence and uniqueness of stochastic differential equations only prove that the solution is locally defined. In the deterministic case, this quadratic property could imply existence of a blow-up. To prove that the solutions do not blow up in finite time in the deterministic case ( $\varepsilon = 0$ ), we can use the fact that the Hamiltonian of the system is constant in time. By adapting that proof (also closely related to the proposition 2) and controlling the Hamiltonian, we will prove that the solutions are defined for all time.

A first remark we will use is the following, for any  $\alpha \in \mathbb{R}^d$  and  $z \in \mathbb{R}^d$ ,

$$(42) \quad \langle \alpha, K_V(z, z) \alpha \rangle_{\mathbb{R}^d} \leq C^2 |\alpha|_{\mathbb{R}^d}^2.$$

Thus, we introduce the stopping times defined as follows: let  $M > 0$  be a constant and

$$(43) \quad \tau_M = \{t \geq 0 \mid \max(|x_t|, |p_t|) \geq M\},$$

let also  $\tau_\infty = \lim_{M \rightarrow \infty} \uparrow \tau_M$  be the explosion time. Differentiating  $H_0(p_{t \wedge \tau_M}, x_{t \wedge \tau_M})$  with respect to  $t$ , we get on ( $t < \tau_M$ ):

$$dH_0(t) = \partial_x H_0(p_t, x_t) dx_t + \partial_p H_0(p_t, x_t) dp_t + \sum_{i=1}^n \text{tr}(K_V(x_i(t), x_i(t))) \frac{\varepsilon^2}{2} dt.$$

In the deterministic case the Hamiltonian is constant, whereas here the stochastic perturbation gives

$$\partial_x H_0(p_t, x_t) dx_t + \partial_p H_0(p_t, x_t) dp_t = \varepsilon \partial_p H_0(p_t, x_t) dB_t.$$

Thus

$$\int_0^{T \wedge \tau_M} dH_0(t) = \int_0^{T \wedge \tau_M} \varepsilon \langle \partial_p H_0(p_t, x_t), dB_t \rangle + \int_0^{T \wedge \tau_M} \sum_{i=1}^n \text{tr}(K_V(x_i(t), x_i(t))) \frac{\varepsilon^2}{2} dt,$$

and

$$(44) \quad E[H_0(p_{T \wedge \tau_M}, x_{T \wedge \tau_M})] \leq H_0(0) + E(C^2 \frac{\varepsilon^2}{2} dn T \wedge \tau_M) \leq H_0(0) + C^2 \varepsilon^2 dn T.$$

<sup>2</sup>Note that we have implemented here a least square approximation algorithm (see (18) and (39) and not an exact interpolation algorithm but with a value  $\gamma$  weighting the data attachment term high enough to make the data error negligible.

Now, we aim at controlling  $x_{t \wedge \tau_M}$  using the control on  $dx_t$  given by  $|\partial_p H_0(p_t, x_t)|_\infty \leq C\sqrt{H_0(p_t, x_t)}$ :

$$(45) \quad |x_{\tau_M \wedge t}| \leq |x_0| + \int_0^{\tau_M \wedge t} CH_0(p_s, x_s)^{1/2} ds \leq |x_0| + \int_0^{\tau_M \wedge t} CH_0(p_{s \wedge \tau_M}, x_{s \wedge \tau_M})^{1/2} ds \\ \leq A_t \doteq |x_0| + \int_0^{\tau_\infty \wedge t} CH_0(p_{s \wedge \tau_\infty}, x_{s \wedge \tau_\infty})^{1/2} ds.$$

However,  $0 \leq A_t$   $P$  a.s. and by monotone convergence theorem (recall that  $H_0$  is non negative),

$$E(A_t) = \lim_{M \rightarrow \infty} (|x_0| + E \left( \int_0^{t \wedge \tau_M} CH_0(p_{s \wedge \tau_M}, x_{s \wedge \tau_M})^{1/2} ds \right)).$$

Also,

$$(46) \quad E \left( \int_0^{t \wedge \tau_M} H_0(p_{s \wedge \tau_M}, x_{s \wedge \tau_M})^{1/2} ds \right) \leq E \left( \int_0^t H_0(p_{s \wedge \tau_M}, x_{s \wedge \tau_M})^{1/2} ds \right) \\ \stackrel{Fub.}{=} \int_0^t E \left( H_0(p_{s \wedge \tau_M}, x_{s \wedge \tau_M})^{1/2} \right) ds \stackrel{Jen.}{\leq} \int_0^t E \left( H_0(p_{s \wedge \tau_M}, x_{s \wedge \tau_M}) \right)^{1/2} ds \\ \stackrel{CS+ (44)}{\leq} \sqrt{t} \left( \int_0^t (H_0(0) + C^2 \varepsilon^2 dn_s) ds \right)^{1/2}.$$

We deduce

$$E(A_t) \leq |x_0| + C\sqrt{t} \left( \int_0^t (H_0(0) + C^2 \varepsilon^2 dn_s) ds \right)^{1/2} < \infty \text{ and } A_t < \infty \text{ } P \text{ a.s.}$$

and as a consequence

$$\limsup_{M \rightarrow \infty} |x_{t \wedge \tau_M}| < +\infty \text{ } P \text{ a.s.}$$

We also control the evolution equation of the momentum as follows,

$$(47) \quad |p_{t \wedge \tau_M}| \leq \int_0^{t \wedge \tau_M} |\partial_x H_0(p_s, x_s)| ds + |p_0| + \int_0^{t \wedge \tau_M} \varepsilon dB_s.$$

Now we use the assumption (3) to control  $\partial_x H_0(p, x)$ :

$$|\partial_x H_0(p, x)| \leq |p| |dv(x)| \leq C|p| H_0^{1/2}.$$

We rewrite inequality (47) and we use Gronwall's Lemma to get:

$$|p_{t \wedge \tau_M}| \leq \int_0^{t \wedge \tau_M} C|p_s| H_0(p_s, x_s)^{1/2} ds + |p_0| + \int_0^{t \wedge \tau_M} \varepsilon dB_s, \\ |p_{t \wedge \tau_M}| \leq \left( |p_0| + \sup_{u \leq t} \left| \int_0^{u \wedge \tau_M} \varepsilon dB_s \right| \right) e^{\int_0^{t \wedge \tau_M} CH_0(p_s, x_s)^{1/2} ds}, \\ |p_{t \wedge \tau_M}| \leq \left( |p_0| + \sup_{u \leq t \wedge \tau_\infty} \left| \int_0^u \varepsilon dB_s \right| \right) e^{\int_0^{t \wedge \tau_\infty} CH_0(p_s, x_s)^{1/2} ds}.$$

The first term on the right-hand side  $|p_0| + \sup_{u \leq t \wedge \tau_\infty} \left| \int_0^u \varepsilon dB_s \right|$  is bounded by  $|p_0| + \sup_{u \leq t} \left| \int_0^u \varepsilon dB_s \right| < \infty$   $P$  a.s. and with inequality (46) we have that

$$e^{\int_0^{t \wedge \tau_\infty} CH_0(p_s, x_s)^{1/2} ds} < \infty \text{ } P \text{ a.s.}$$

Since on  $(\tau_\infty \leq t)$  one has

$$\lim_{M \rightarrow \infty} \max(|x_{t \wedge \tau_M}|, |p_{t \wedge \tau_M}|) = \lim_{M \rightarrow \infty} |p_t| = \infty,$$

we deduce  $P(\tau_\infty \leq t) = 0$  and  $\tau_\infty = +\infty$  almost surely.

We have proved for the case  $\varepsilon(p, x) = \varepsilon Id$ ,

**Theorem 4.** *Under assumption (3), the solutions of the stochastic differential equation defined by*

$$\begin{aligned} dp_t &= -\partial_x H_0(p_t, x_t)dt + \varepsilon(p_t, x_t)dB_t \\ dx_t &= \partial_p H_0(p_t, x_t)dt. \end{aligned}$$

*are non exploding when  $\varepsilon : \mathbb{R}^{nd} \times \mathbb{R}^{nd} \mapsto L(\mathbb{R}^{nd})$  is a Lipschitz and bounded map.*

*Proof.* To extend the proof to the case when  $\varepsilon$  is a Lipschitz and bounded map of  $p$  and  $x$ , we can prove that the preceding inequalities are still valid.

First, with the Lipschitz property of  $\varepsilon$  the solutions are still defined locally. The Ito formula is now written as, on  $(t < \tau_M)$

$$dH_0(t) = \partial_x H_0(p_t, x_t)dx_t + \partial_p H_0(p_t, x_t)dp_t + \frac{1}{2}\text{tr}(\varepsilon^T(p_t, x_t)K_{x_t}\varepsilon(p_t, x_t))dt.$$

where  $K_x$  is block matrix defined by  $K_x \doteq (K_V(x_i, x_j))_{1 \leq i, j \leq n}$ .

We still have the inequality (44) with

$$\text{tr}(\varepsilon^T(p_t, x_t)K_{x_t}\varepsilon(p_t, x_t)) \leq (Cnd|\varepsilon|_\infty)^2$$

if  $|\varepsilon(p, x)w|^2 \leq |\varepsilon|_\infty^2 |w|_\infty^2$  where  $|\varepsilon|_\infty$  denotes the supremum norm. Indeed, if  $(e_i)_{i \in [1, nd]}$  the canonical basis of  $\mathbb{R}^{nd}$ , denoting  $\varepsilon \doteq \varepsilon(x, p)$ , we have

$$\text{tr}(\varepsilon^t K_x \varepsilon) = \sum_{i=1}^{nd} \langle \varepsilon(e_i), K_x \varepsilon(e_i) \rangle \leq \lambda_{K_x}^* \sum_{i=1}^{nd} \langle \varepsilon(e_i), \varepsilon(e_i) \rangle,$$

where  $\lambda_{K_x}^*$  is the largest eigenvalue of  $K_x$ . We have  $\lambda_{K_x}^* \leq \text{tr}(K_x) = \sum_{i=1}^n \text{tr}(K_V(x_i, x_i))$  and using (42) we get  $\text{tr}(K_V(x_i, x_i)) \leq d\lambda_{K_V(x_i, x_i)}^* \leq dC^2$  so that  $\lambda_{K_x}^* \leq C^2 nd$ . Hence,

$$\text{tr}(\varepsilon^t K_x \varepsilon) \leq C^2 nd \sum_{i=1}^{nd} |\varepsilon|_\infty^2 \leq (Cnd|\varepsilon|_\infty)^2.$$

Thus we get,

$$\begin{aligned} \int_0^{T \wedge \tau_M} dH_0(t) &\leq \int_0^{T \wedge \tau_M} \langle \partial_p H_0(p_t, x_t), \varepsilon(p_t, x_t)dB_t \rangle + \int_0^{T \wedge \tau_M} \frac{(Cnd|\varepsilon|_\infty)^2}{2} dt, \\ E[H_0(T \wedge \tau_M)] &\leq H_0(0) + E\left(\frac{(Cnd|\varepsilon|_\infty)^2}{2} T \wedge \tau_M\right) \leq H_0(0) + (Cnd|\varepsilon|_\infty)^2 T, \end{aligned}$$

and all the remaining inequalities follow easily thanks to the control on  $H_0$  and the bound on  $\varepsilon$ .  $\square$

Once this stochastic model is well-posed on landmark space, the question of its extension to shape spaces naturally arises. It can be proved that this stochastic model does have an extension to the infinite dimensional case: in the case of  $L^2(\mathbb{R}/\mathbb{Z}, \mathbb{R}^2)$ , the natural extension of the Brownian motion on the landmark space is a cylindrical Brownian motion on  $L^2(\mathbb{R}/\mathbb{Z}, \mathbb{R}^2)$ . It may be somewhat surprising to deal with such irregular noise on the momentum variable, we stress the fact that this noise is read by the kernel which strongly regularizes the noise. Though we will not develop it further, it proves that this model has a consistent extension to continuous-shape spaces and could be used to deal with random evolutions of continuous shapes.

**6.2. Simulations.** With these simulations we illustrate the interesting features we observed above. First, this model gives realistic perturbations of geodesics contrary to a first order model such as a Kunita flow. A simulation of a Kunita flow is illustrated in figure Fig. 7 where the evolution of 40 points on the unit circle is represented under a Gaussian kernel of width 0.9. The time evolution has, as expected, the roughness of a Brownian motion and the space variation is smoother due to the kernel. In comparison, our stochastic model gives smoother evolutions in time as in Fig. 10 and Fig. 11.

Second, our stochastic model is a perturbation of a geodesic evolution and this nice property is illustrated in figures Fig. 9 - 11.

Figure Fig. 8 shows the geodesic evolution of 40 equidistributed points on the unit circle for a Gaussian kernel of width 1.0, the target configuration for the landmarks is obtained through a simple affine transformation that gives the final ellipse. On these

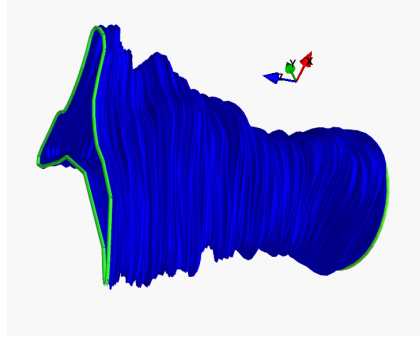


FIGURE 7. A simulation of Kunita flow with 40 points on the unit circle on the left of the figure. The  $z$  axis (blue arrow) represents the time.

simulations the color change only represents time. Figures Fig. 9 - 11 represent stochastic perturbations of the previous geodesic; we progressively increase the standard deviation  $\epsilon$  of the noise from  $\sqrt{n}\epsilon = 0.9$  to 1.7 and finally  $\sqrt{n}\epsilon = 3.5$  (the noise is rescaled w.r.t. the number of landmarks to converge to a well defined SPDE at the limit (see [34])). Each of these three figures represents one Monte-Carlo simulation of the stochastic model with a simple Euler scheme.

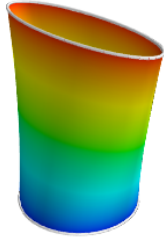


FIGURE 8. Geodesic evolution  
- White unit circle as initial shape.

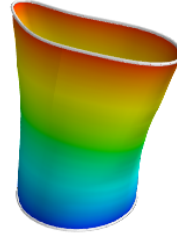


FIGURE 9. White noise perturbation of the geodesic (same initial momentum  $p_0$ ),  $\sqrt{n}\epsilon = 0.9$

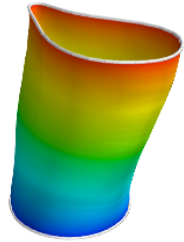


FIGURE 10. Increasing the variance of the noise,  $\sqrt{n}\epsilon = 1.7$

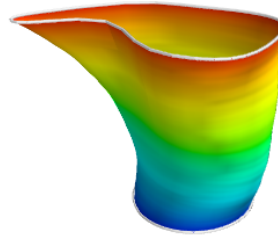


FIGURE 11. Increasing the variance of the noise,  $\sqrt{n}\epsilon = 3.5$

The simulations in figures Fig. 12 show the position at time 1 of the 40 points for 5 Monte-Carlo simulations. On the two figures we plotted the initial momentum  $p_0$  (attached to the 40 points) associated with the geodesic from the initial circle to the target ellipse. As this model was designed to produce random shape evolutions, it can also be used as a generative engine to produce random shapes. Increasing the noise also increases the expectation of the energy of the system since the Ito formula applied on the Hamiltonian in subsection 6.1 shows a linear growth in time of  $H_0$  proportional to  $\epsilon$ . This can

be guessed when comparing the two displayed cases in Fig. 12 since in the second one the noise is 4 times bigger. For any statistical estimation of the model parameters, this property should be somehow taken into account.

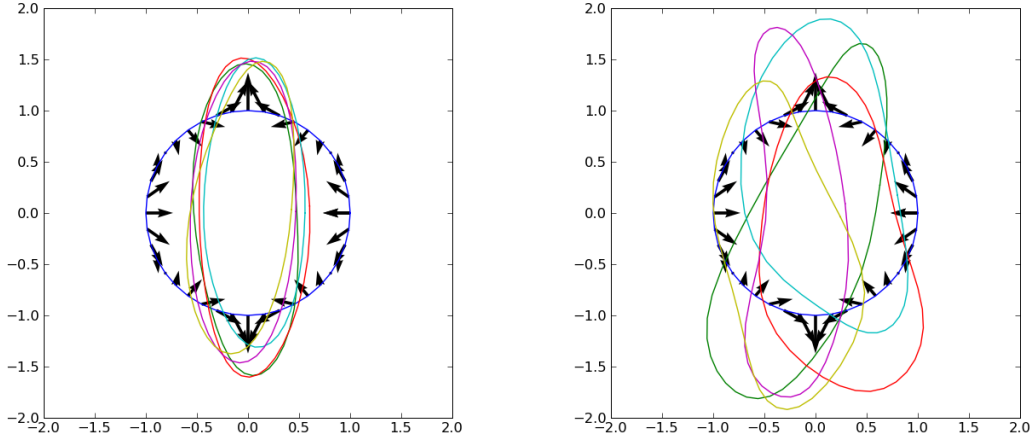


FIGURE 12. Effects of the standard deviation  $\epsilon$  of the noise : 5 simulations of random deformations of the unit circle. The initial momentum  $p_0$  is fixed, the kernel width  $\lambda = 1$  (see (4)). Left-hand side :  $\sqrt{n}\epsilon = 0.25$  ; right-hand side :  $\sqrt{n}\epsilon = 1$ .

An important feature of this model is that the noise is "read" by the kernel. We show in Fig. 13 simulations of the model for a null initial momentum on the same initial shape and we decrease the width of the Gaussian kernel from 3 to 0.3. The standard deviation of the noise is constant set to  $\sqrt{n}\epsilon = 1.0$ .

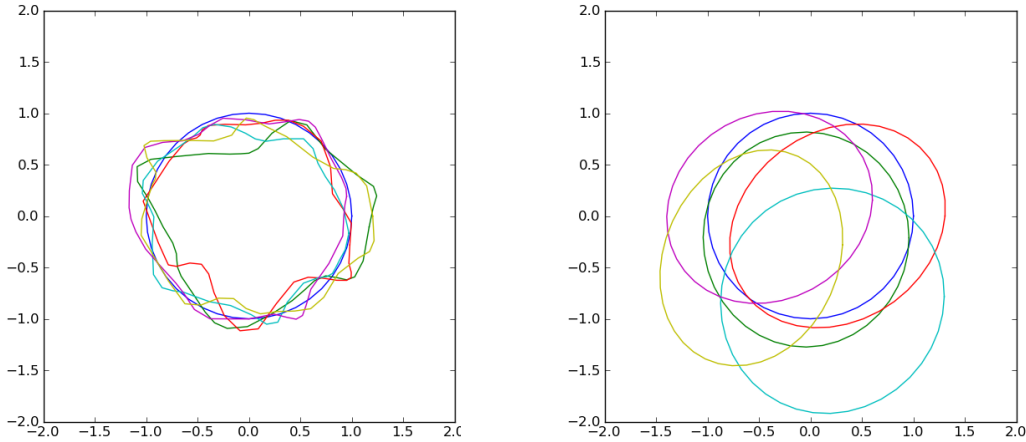


FIGURE 13. Effects of the width of the Gaussian kernel. Left-hand side :  $\lambda = 0.3$  ; right-hand side :  $\lambda = 3.0$ .

These last simulations show the importance of the choice of the kernel and as a by-product the choice of the operator  $\varepsilon$  in front of the noise will be also important. Now we can formulate a stochastic model for evolutions of shape that would be closer to realistic evolutions:

$$(48) \quad \begin{cases} dp_t = -\partial_q H_0(p_t, q_t) + u_t + \varepsilon dB_t \\ dq_t = \partial_p H_0(p_t, q_t), \end{cases}$$

where  $u_t$  is of bounded variations. At this point, we underline that the model parameterization is completely open and it should be tightly related to consistent statistical estimations.

## 7. SHAPE SPLINES ON HOMOGENEOUS SPACE

In this section we provide a more formal and geometrical picture of what could be an extension of the shape spline to the previously mentioned important cases. For this, we need to introduce some of the standard vocabulary of geometrical mechanics as developed in [25]. We will try as much as possible to avoid the conceptual burden of the intrinsic differential and symplectic geometry through the extensive use of local coordinates. Readers looking for a more intrinsic formulation could refer to [25].

**7.1. Geometrical setting.** The proposed framework for shape spline is given by three ingredients: a group  $G$  of transformations, a Riemannian manifold  $Q$  and a *left* action of  $G$  on  $Q$  denoted  $(g, q) \rightarrow L_q(g) \doteq g \cdot q$ . We will assume also that for any  $q \in Q$ ,  $g \rightarrow g \cdot q$  is a surjective submersion (i.e. the differential of  $L_q$  has full rank everywhere).

**7.1.1. Local coordinates.** Basically  $G$  will be a (finite dimensional) Lie group with Lie algebra  $\mathfrak{G}$  on which we consider a right invariant metric given by a dot product on  $\mathfrak{G}$ . The infinite dimensional setting where  $G$  is a group of diffeomorphisms is more involved and requires more analytical work as in [33]. This is clearly the target setting we have in mind but we want in this paper to stay away from any complicated analytical developments. To keep the focus on the global picture, we will assume implicitly that we work in a finite dimensional setting for which the existence of all the introduced objects is straightforward.

Denoting  $q = (q^1, \dots, q^n)$  local coordinates on  $Q$  and since  $(dq^1, \dots, dq^n)$  is a basis of  $T_q^*Q$ , we can write any  $\alpha \in T_q^*Q$  as  $\alpha = \sum p_i dq^i$  so that  $(q^1, \dots, q^n, p_1, \dots, p_n)$  are local coordinates on the cotangent bundle  $T^*Q$ . Given  $q = (q^1, \dots, q^n)$ , we will denote  $p = (p_1, \dots, p_n)$  a generic element of  $T_q^*Q$  and  $m = (q, p)$  a generic element of  $T^*Q$  as we did previously in the flat case of landmarks.

**7.1.2. Infinitesimal actions and cotangent lift.** The first thing we need is to extend the action of  $G$  on  $Q$  to an action of  $G$  on the cotangent bundle  $T^*Q$ . Note that the differentiation in  $g$  of  $g \rightarrow q \cdot q$  at  $g = \text{Id}_G$  yields an infinitesimal action  $(\xi, q) \rightarrow \xi \cdot q$  for  $\xi \in \mathfrak{G}$ . Differentiation in  $q$  yields the action  $(g, \delta q) \rightarrow g \cdot \delta q \in T_{g \cdot q}Q$  for  $\delta q \in T_qQ$  and by duality the action  $(g, p) \rightarrow g \cdot p \in T_{g \cdot q}^*Q$  for  $p \in T_q^*Q$ , uniquely defined through the equality

$$(49) \quad (g \cdot p | g \cdot \delta q) \doteq (p | \delta q).$$

We denote

$$(50) \quad (g, m) \rightarrow m \cdot g \doteq (g \cdot q, g \cdot p)$$

for  $m = (q, p) \in T^*Q$  the induced action on  $T^*Q$ .

In summary, the initial action  $g \rightarrow g \cdot q$  on  $Q$  is naturally lifted to an action  $g \rightarrow g \cdot m$  on the cotangent space  $T^*Q$  (usually called *cotangent lift* [25]). Differentiating the action  $(g, m) \rightarrow g \cdot m$  on the cotangent space  $T^*Q$  at  $g = \text{Id}_G$ , we get the infinitesimal action on  $T^*Q$ , which is defined in local coordinates by  $\xi \cdot m = (\xi \cdot q, \xi \cdot p)$  where

$$(51) \quad (\xi \cdot p | \delta q) + (p | \xi \cdot \delta q) = 0$$

as obtained by differentiation of the conservation equation (49).

**7.2. Euler-Poincaré equation.** The initial matching problem between shapes in  $Q$  is defined as the solution of the optimal control problem with fixed boundary

$$\left| \begin{array}{l} \min_{\xi_t} \frac{1}{2} \int_0^1 (L\xi_t, \xi_t) dt \\ \text{subject to} \\ \dot{q}_t = \xi_t \cdot q_t, \quad q_0 = q_{\text{init}}, \quad q_1 = q_{\text{targ}} \end{array} \right.$$

where  $L : \mathfrak{G}^* \rightarrow \mathfrak{G}$  is the isometry between  $\mathfrak{G}$  and its dual induced by the metric on  $\mathfrak{G}$ .

The Hamiltonian associated to the classical matching problem is given in local coordinates by  $H(q, p, \xi) = (p | \xi \cdot q) - \frac{1}{2} (L\xi | \xi)$  with reduced form

$$H(q, p) \doteq \frac{1}{2} (KJ(q, p) | J(q, p))$$

where  $J(q, p) \in \mathfrak{G}^*$  is uniquely defined by

$$(52) \quad (J(q, p) | \xi) = (p | \xi \cdot q)$$

for any  $\xi \in \mathfrak{G}$  (usually called the *momentum map*). To obtain the Hamiltonian evolution, we need to compute the variation of  $\delta J$  as a function of the variation  $\delta q$  and  $\delta p$  in  $q$  and  $p$ . Introducing in local coordinates the so-called *symplectic matrix*

$$\mathbb{J} \doteq \begin{pmatrix} 0 & \text{Id}_n \\ -\text{Id}_n & 0 \end{pmatrix}$$

one checks easily that  $(\delta J|\xi) = (p|\xi \cdot \delta q) + (\delta p|\xi \cdot q)$  so that using (51) we get

$$(53) \quad (\delta J|\xi) = -(\xi \cdot p|\delta q) + (\delta p|\xi \cdot q) = -(\mathbb{J}(\xi \cdot m)|\delta m)$$

and  $dH = -\mathbb{J}(KJ \cdot m)$ . Since the associated Hamiltonian evolution equation are given by  $\dot{m} = \mathbb{J}dH$  we get

$$(54) \quad \dot{m} = Kj \cdot m \text{ with } j = J(q, p)$$

or equivalently

$$(55) \quad \dot{q} = \xi \cdot q \text{ and } \dot{p} = \xi \cdot p \text{ with } \xi = KJ(q, p).$$

This extremely simple expression of  $\dot{m}$  in term of the momentum map and the infinitesimal action on the cotangent space reveals part of the nice geometrical structure underlying the evolution. In this setting, it is interesting to consider the time evolution of the pair  $(q, j)$  instead of the pair  $(q, p)$  since  $j$  follows an autonomous equation. Indeed, from (51) and (55), we get that for any  $\zeta \in \mathfrak{G}$  we have  $(\frac{dj}{dt} | \zeta) = (\xi_t \cdot p | \zeta \cdot q) + (p | \zeta \cdot (\xi_t \cdot q)) = (p | \zeta \cdot (\xi_t \cdot q) - \xi_t \cdot (\zeta \cdot q)) = -(p | \text{ad}_{\xi_t}(\zeta) \cdot q) = -(j | \text{ad}_{\xi_t}(\zeta))$  where  $\text{ad}_{\xi}(\zeta) = [\xi, \zeta]$  is the adjoint representation of the Lie algebra  $\mathfrak{G}$  so that we get

$$\frac{dj}{dt} + \text{ad}_{Kj}^* j = 0.$$

This equation, called the *Euler-Poincaré equation* plays a central role in geometric mechanics and more recently in the large deformation methods in shape analysis and computational anatomy [16, 25].

Now consider the perturbed dynamic  $\dot{p} = Kj(m).p + u$  with  $u \in T_q^*Q$  or equivalently

$$\frac{dj}{dt} + \text{ad}_{Kj}^* j = h$$

where  $h \doteq J(q, u)$  and the associated optimal control problem for the state variables  $(q, j) \in Q \times \mathfrak{G}^*$  and the cost  $\frac{1}{2}|u|_{q,*}^2$ . The norm  $|u|_{q,*}$  we consider here is the dual norm induced on  $T_q^*Q$  by the metric on  $T_qQ$ . Let  $\mathcal{K}_q : T_q^*Q \rightarrow T_qQ$  be the isometry such that  $(u | \mathcal{K}_q u) = |u|_{q,*}^2$ . The Hamiltonian associated to our new control problem for the costate variable  $(p_q, \zeta) \in T_q^*Q \times \mathfrak{G}$  is given by

$$(56) \quad \mathcal{H}(q, j, p_q, \zeta, u) = (p_q | Kj \cdot q) + (-\text{ad}_{Kj}^* j + h | \zeta) - \frac{1}{2}(\mathcal{K}_q u | u).$$

To compute its reduced form let us note that from (51), we get  $\frac{\partial}{\partial u}(h | \zeta) = \zeta \cdot q$ . Hence  $\frac{\partial}{\partial u}\mathcal{H} = 0$  implies  $\mathcal{K}_q u = \zeta \cdot q$  giving the reduced Hamiltonian

$$(57) \quad \mathcal{H}(q, j, p_q, \zeta) = (p_q | Kj \cdot q) + (-\text{ad}_{Kj}^* j | \zeta) + \frac{1}{2}|\zeta \cdot q|_q^2$$

$$(58) \quad = (J(q, p_q) | Kj) + (j | \text{ad}_{\zeta}(Kj)) + \frac{1}{2}|\zeta \cdot q|_q^2,$$

where  $| \cdot |_q$  denote the norm on  $T_qQ$  given by the Riemmanian metric on  $Q$ . The associated Hamiltonian evolution is derived quite easily:

$$(59) \quad \begin{cases} \xi = Kj \\ \frac{dq}{dt} = \xi \cdot q \\ \frac{dp_q}{dt} - \xi \cdot p_q = -\frac{\partial}{\partial q}(\frac{1}{2}|\zeta \cdot q|_q^2) \\ \frac{dj}{dt} + \text{ad}_{\xi}^*(j) = \frac{\partial}{\partial \zeta}(\frac{1}{2}|\zeta \cdot q|_q^2) \\ \frac{d\zeta}{dt} + \text{ad}_{\zeta}(\xi) + K\text{ad}_{\zeta}^*(j) + KJ(q, p_q) = 0 \end{cases}$$

Here again, the derivation should be considered at a formal level or in a smooth finite dimensional setting since existence results are beyond the scope of this paper.

## 8. CONCLUSION

In this paper, we present new tools for shape evolution analysis or growth analysis in computational anatomy through the introduction of second order evolutions. Shape splines seem to overcome some of the limitations of the previous first order schemes and with their stochastic counterpart can provide the backbone of new statistical time regression tools. From various perspectives, shape splines offer new interesting mathematical and practical challenges: extensions to the infinite dimensional case of continuous shapes and to images, development of efficient and scalable numerical schemes to solve the spline estimation problem, integration of time realignment as developed in [9], derivation of more complex stochastic engines beyond the white-noise situation presented here and development of consistent statistical schemes in the spirit of [3]. The solutions to some of these problems appear well within reach.

In this paper we preferred to stick to the finite dimensional setting in order to focus on the general picture and avoid more difficult analysis, the more technically involved infinite dimensional situation has been partially explored in the stochastic case in [34].

**Acknowledgments.** The authors would like to thank Darryl D. Holm and Colin J.

Cotter for a fruitful discussion about introducing noise on the momentum variable which was the starting point of the stochastic model presented here.

## 9. APPENDIX

**9.1. Link with cubic splines.** We recall that on a Riemannian manifold  $(M, g_M)$ , a cubic spline between  $(x_0, v_0)$  and  $(x_1, v_1)$  is a  $C^2$  curve  $c : [0, 1] \rightarrow M$  that minimizes

$$(60) \quad \mathcal{J}(c) = \frac{1}{2} \int_0^1 g_M(\nabla_{\dot{c}} \dot{c}, \nabla_{\dot{c}} \dot{c}) dt.$$

The Euler-Lagrange equation for this functional is the following (see [29, 11])

$$(61) \quad \nabla_t^3 \dot{c} + R(\nabla_t \dot{c}, \dot{c}) \dot{c} = 0.$$

Let  $c : I \rightarrow M$  a smooth path, then the covariant derivative is given in coordinates by

$$\nabla_{\dot{c}} \dot{c} = [\ddot{c}_i + \sum_{k,l} \dot{c}_k \Gamma_{k,l}^i(c) \dot{c}_l]_{i \in [1,r]},$$

and we observe that the second member of the right hand side only depends on  $(\dot{c}, c)$ . We compare this expression with the Hamiltonian equations (here  $M = \mathcal{L}$ ):

$$(62) \quad \ddot{c} = \frac{d}{dt} k(c, c) p = [\partial_1 k(c, c)](\dot{c}) p + k(c, c) \dot{p},$$

$$(63) \quad \ddot{c} = [\partial_1 k(c, c)](k(c, c)^{-1} p) p - k(c, c) \partial_x H + k(c, c) u,$$

where for any  $x, y \in \mathbb{R}^{nd}$ ,  $k(x, y)$  denotes the block matrix defined by

$$k(x, y) \doteq (K_V(x_i, y_j))_{1 \leq i, j \leq n}.$$

Now the geodesic equations are given by

$$\ddot{c}_i + \sum_{k,l} \dot{c}_k \Gamma_{k,l}^i(c) \dot{c}_l = 0 \text{ for } i \in [1, r],$$

where  $\Gamma_{k,l}^i = g_{\mathcal{L}}(\nabla_{\partial_i} \partial_j, \partial_k)$  are the Christoffel symbols in the chosen coordinates.

However, if  $u = 0$  we obtain the geodesic equations in the Hamiltonian form. Then we can identify

$$-[\sum_{k,l} \dot{c}_k \Gamma_{k,l}^i(c) \dot{c}_l]_{i \in [1,r]} = [\partial_1 k(c, c)](k(c, c)^{-1} p) p - k(c, c) \partial_x H,$$

which gives

$$\nabla_{\dot{c}} \dot{c} = k(c, c) u.$$

Now we have

$$g_{\mathcal{L}}(\nabla_{\dot{c}} \dot{c}, \nabla_{\dot{c}} \dot{c}) = g_{\mathcal{L}}(k(c, c) u, k(c, c) u) = \langle u, k(c, c) u \rangle,$$

which proves the desired result.

**9.2. Proof of Proposition 1.** Let us denote  $\partial_{12}K_W(z_1, z_2) \doteq \frac{\partial^2 K_W}{\partial z_1 \partial z_2}(z_1, z_2)$  for any  $z_1, z_2 \in \mathbb{R}^d$  and  $\Delta_{z,h,\epsilon} \doteq \frac{1}{\epsilon}(\delta_{z+\epsilon h} - \delta_z)$  for any  $z, h \in \mathbb{R}^d$  and  $\epsilon > 0$ . Let us recall that for any  $z_1, z_2 \in \mathbb{R}^d$

$$(64) \quad \langle \delta_{z_1}, \delta_{z_2} \rangle_{W^*} = K_W(z_1, z_2).$$

First, we start with the proof that  $\Delta_{z,h,\epsilon}$  converges in  $W^*$  when  $\epsilon \rightarrow 0$ . Indeed we get from (64) that

$$(65) \quad \langle \Delta_{z_1, h_1, \epsilon_1}, \Delta_{z_2, h_2, \epsilon_2} \rangle_{W^*} = \int_0^1 \int_0^1 \partial_{12}K_W(z_1 + s\epsilon_1 h_1, z_2 + t\epsilon_2 h_2) \cdot h_1 \otimes h_2 \, ds dt$$

$$(66) \quad = \partial_{12}K_W(z_1, z_2) \cdot h_1 \otimes h_2 + o((|\epsilon_1 h_1| + |\epsilon_2 h_2|)|h_1||h_2|),$$

so that  $\|\Delta_{z,h,\epsilon} - \Delta_{z,h,\epsilon'}\|_{W^*}^2 = o(|\epsilon| + |\epsilon'|)$ . Since  $W^*$  is complete,  $\Delta_{z,h,\epsilon}$  converges in  $W^*$  to a limit point denoted  $\delta'_{z,h}$  such that  $\langle \delta'_{z_1, h_1}, \delta'_{z_2, h_2} \rangle_{W^*} = \lim \langle \Delta_{z_1, h_1, \epsilon}, \Delta_{z_2, h_2, \epsilon} \rangle_{W^*} = \partial_{12}K_W(z_1, z_2) \cdot h_1 \otimes h_2$ . In particular  $h \rightarrow \delta'_{z,h}$  is a continuous linear mapping from  $\mathbb{R}^{nd}$  to  $W^*$ .

Now, we get from (65)

$$\|\delta_{z,h} - (\delta_z + \delta'_{z,h})\|_{W^*}^2 = o(|h|^3)$$

so that  $\psi$  is differentiable at any location  $x \in \mathbb{R}^{nd}$  with differential  $\psi'(x)$  given by

$$\psi'(x)\Delta x \doteq \frac{1}{n} \sum_{i=1}^n \delta'_{x_i, \Delta x_i}.$$

Thus, using again (65), we get  $\langle \psi'(x)\Delta x, \psi'(x)\Delta x \rangle_{W^*} = \frac{1}{n^2} \sum_{i,j=1}^n \partial_{12}K(x_i, x_j) \cdot \Delta x_i \otimes \Delta x_j$  which proves Proposition 1.

## REFERENCES

- [1] A. A. Agrachev and Y. L. Sachkov. *Control theory from the geometric viewpoint*, volume 87 of *Encyclopaedia of Mathematical Sciences*. Springer-Verlag, Berlin, 2004. , Control Theory and Optimization, II.
- [2] J. Ahlberg, E. Nilson, and J. Walsh. The theory of splines and their applications. *Mathematics in Science and Engineering*, 38, 1967.
- [3] S. Allasonnière, Y. Amit, and A. Trouvé. Towards a coherent statistical framework for dense deformable template estimation. *J. R. Statist. Soc. B*, 69(1):3–29, 2007.
- [4] S. Allasonnière, A. Trouvé, and L. Younes. Geodesic shooting and diffeomorphic matching via textured meshes. In *EMMCVPR05*, pages 365–381, 2005.
- [5] M. Camarinha, F. S. Leite, and P. Crouch. Splines of class  $\mathcal{C}^k$  on non-euclidean spaces. *IMA Journal of Mathematical Control & Information*, 12:399–410, 1995.
- [6] P. Crouch and F. S. Leite. The dynamic interpolation problem: On Riemannian manifold, Lie groups and symmetric spaces. *Journal of dynamical & Control Systems*, 1:177–202, 1995.
- [7] B. Davis, P. Fletcher, E. Bullitt, and S. Joshi. Population shape regression from random design data. In *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*, pages 1–7, Oct. 2007.
- [8] I. L. Dryden and K. V. Mardia. *Statistical Shape Analysis*. 1998.
- [9] S. Durrleman, X. Pennec, G. Gerig, A. Trouvé, and N. Ayache. Spatiotemporal atlas estimation for developmental delay detection in longitudinal datasets. In *Medical Image Computing and Computer Assisted Intervention*, September 2009.
- [10] T. Edwin. Jaynes. Information theory and statistical mechanics. *Physical Review*, 106(4):620–630, 1957.
- [11] R. Giambo and F. Giannoni. An analytical theory for riemannian cubic polynomials. *IMA Journal of Mathematical Control & Information*, 19:445–460, 2002.
- [12] J. Glaunes, A. Trouve, and L. Younes. Diffeomorphic matching of distributions: A new approach for unlabelled point-sets and sub-manifolds matching. In *Computer Vision and Pattern Recognition*, volume 2, 2004.
- [13] J. Glaunès, A. Trouvé, and L. Younes. Modeling planar shape variation via Hamiltonian flows of curves. In H. Krim and A. Yezzi, editors, *Statistics and Analysis of Shapes*, pages 335–361. Springer Birkhauser, 2006.
- [14] U. Grenander, A. Srivastava, and S. Saini. Characterization of biological growth using iterated diffeomorphisms. In *ISBI*, pages 1136–1139, 2006.
- [15] U. Grenander, A. Srivastava, and S. Saini. A pattern-theoretic characterization of biological growth. *IEEE Trans. Med. Imaging*, 26(5):648–659, 2007.
- [16] D. R. Holm, A. Trouvé, and L. Younes. The Euler Poincaré theory of metamorphosis. *Quarterly of Applied Mathematics*, 2009. (to appear).
- [17] R. V. Iyer, R. Holsapple, and D. Doman. Optimal control problems on parallelizable Riemannian manifolds: Theory and applications. *ESAIM. COCV*, 12:1–11, 2006.

- [18] J. Jackson. *Dynamic interpolation and application to flight control*. PhD thesis, Arizona State University, 1990.
- [19] S. Joshi and M. Miller. Landmark matching via large deformation diffeomorphisms. *International Journal of Computer Vision*, 2000.
- [20] J. Kapur. *Maximum-entropy models in science and engineering*. Wiley-Interscience, 1989.
- [21] D. G. Kendall. The diffusion of shape. *Advances in Applied Probability*, vol 9:pp. 428–430, 1977.
- [22] A. Khan and M. Beg. Representation of time-varying shapes in the large deformation diffeomorphic framework. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 1521–1524, May 2008.
- [23] H. Kunita. *Stochastic flows and Stochastic Differential Equations*. Cambridge Studies in Advanced Mathematics, 1997.
- [24] J. Macki and A. Strauss. *Introduction to optimal control theory*. Springer, 1982.
- [25] J. Marsden, T. Ratiu, and Maïsser. *Introduction to mechanics and symmetry*. Springer, 1999.
- [26] P. W. Michor and D. Mumford. An overview of the riemannian metrics on spaces of curves using the hamiltonian approach. *Applied and Computational Harmonic Analysis*, 23:74, 2007.
- [27] M. I. Miller, A. Trouvé, and L. Younes. On the metrics and Euler-Lagrange equations of computational anatomy. *Annual Review of biomedical Engineering*, 4:375–405, 2002.
- [28] M. I. Miller, A. Trouvé, and L. Younes. Geodesic shooting for computational anatomy. *J. Math. Imaging Vis.*, 24(2):209–228, 2006.
- [29] L. Noakes, G. Heinzinger, and B. Paden. Cubic splines on curved spaces. *IMA Journal of Mathematical Control & Information*, 6:465–473, 1989.
- [30] N. Portman, U. Grenander, and E. R. Vrsay. Direct estimation of biological growth properties from image data using the "grid" model. In *ICIAI*, pages 832–843, 2009.
- [31] I. Schoenberg. Contributions to the problem of approximation of equidistant data by analytic functions. *Quart. Appl. Math.* 4, 45-99 (Part A), 112-141 (Part B), 1946.
- [32] A. Srivastava, S. Saini, Z. Ding, and U. Grenander. Maximum-likelihood estimation of biological growth variables. pages 107–118, 2005.
- [33] A. Trouvé and L. Younes. Local geometry of deformable templates. *Siam Journal of Mathematical Analysis*, 2005.
- [34] F.-X. Vialard. *Hamiltonian Approach to Shape Spaces in a Diffeomorphic Framework : From the Discontinuous Image Matching Problem to a Stochastic Growth Model*. PhD thesis, ENS Cachan, 2009.
- [35] L. Younes, F. Arrate, and M. I. Miller. Evolutions equations in computational anatomy. *NeuroImage*, 45(1, Supplement 1):S40 – S50, 2009. Mathematics in Brain Imaging.

(Alain Trouvé) CMLA, ECOLE NORMALE SUPÉRIEURE DE CACHAN, CNRS, UNIVER SUD, 61, AVENUE DU PRÉSIDENT WILSON, F-94 235 CACHAN CEDEX  
*E-mail address:* alain.trouve@cmla.ens-cachan.fr

(François-Xavier Vialard) INSTITUTE FOR MATHEMATICAL SCIENCE, IMPERIAL COLLEGE LONDON, 53 PRINCE'S GATE, SW7 2PG, LONDON, UK  
*E-mail address:* francois.xavier.vialard@normalesup.org